UNITED STATES PATENT AND TRADEMARK OFFICE
DOCUMENT CLASSIFICATION BARCODE SHEET

000

CATEGORY:

CLEARED

ADDRESS
CONTACT IF FOUND:

404 Rec'd PCT/PTO 08 SEP 1998

FORM PTO-1390 (Modified) (REV 10-95) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: **1038-833 MIS** PCT

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR: 09/142628

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/CA97/00163	7 March 1997	8 March 1996

TITLE OF INVENTION
TRANSFERRIN RECEPTOR GENES OF MORAXELLA

APPLICANT(S) FOR DO/EO/US
Lisa E. Myers, et al.

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☐ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
   a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☒ has been transmitted by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☐ A copy of the International Search Report (PCT/ISA/210).
8. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ have been transmitted by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☐ have not been made and will not be made.
9. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
10. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)). - unsigned copy
11. ☐ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).

Items 13 to 18 below concern document(s) or information included:

13. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
15. ☒ A **FIRST** preliminary amendment.
    A **SECOND** or **SUBSEQUENT** preliminary amendment.
16. ☐ A substitute specification.
17. ☐ A change of power of attorney and/or address letter.
18. ☐ Certificate of Mailing by Express Mail
19. ☐ Other items or information:

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR	INTERNATIONAL APPLICATION NO. PCT/CA97/00163	ATTORNEY'S DOCKET NUMBER 1038-833 MIS

20. The following fees are submitted:.

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

		CALCULATIONS	PTO USE ONLY
☐ Search Report has been prepared by the EPO or JPO	**$930.00**		
☐ International preliminary examination fee paid to USPTO (37 CFR 1.482)	**$720.00**		
☐ No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2))	**$790.00**		
☒ Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2) paid to USPTO	**$1,070.00**		
☐ International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4)	**$98.00**		
ENTER APPROPRIATE BASIC FEE AMOUNT =		**$1,070.00**	
Surcharge of **$130.00** for furnishing the oath or declaration later than months from the earliest claimed priority date (37 CFR 1.492 (e)).	☐ 20 ☐ 30	**$0.00**	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	25 - 20 =	5	x $22.00	**$110.00**	
Independent claims	10 - 3 =	7	x $82.00	**$574.00**	
Multiple Dependent Claims **(check if applicable).**			☐	**$0.00**	
	TOTAL OF ABOVE CALCULATIONS		=	**$1,754.00**	
Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28) **(check if applicable).**			☐	**$0.00**	
	SUBTOTAL		=	**$1,754.00**	
Processing fee of **$130.00** for furnishing the English translation later than months from the earliest claimed priority date (37 CFR 1.492 (f)).	☐ 20 ☐ 30		+	**$0.00**	
	TOTAL NATIONAL FEE		=	**$1,754.00**	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) **(check if applicable).**			☐	**$0.00**	
	TOTAL FEES ENCLOSED		=	**$1,754.00**	
				Amount to be: refunded	$
				charged	$

☒ A check in the amount of **$1,754.00** to cover the above fees is enclosed.

☐ Please charge my Deposit Account No. in the amount of to cover the above fees.
A duplicate copy of this sheet is enclosed.

☒ The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. **19-2253** A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Michael I. Stewart
Sim & McBurney
6th Floor, 330 University Avenue
Toronto, Ontario
Canada, M5G 1R7.

The PTO did not receive the following listed item(s) a check of $1,754.00

SIGNATURE

Michael I. Stewart
NAME

24,973
REGISTRATION NUMBER

September 3, 1998
DATE

62 Rec'd PCT/PTO 08 SEP 1998

09/142628

TITLE OF INVENTION

TRANSFERRIN RECEPTOR GENES OF MORAXELLA

FIELD OF INVENTION

The present invention relates to the molecular cloning of genes encoding transferrin receptor (TfR) proteins and, in particular, to the cloning of transferrin receptor genes from *Moraxella (Branhamella) catarrhalis*.

REFERENCE TO RELATED APPLICATION

This application is a continuation-in-part of copending United States Patent Application No. 08/778,570 filed January 3, 1997, which itself is a continuation-in-part of United States Patent Application No. 08/613,009 filed March 8, 1996.

BACKGROUND OF THE INVENTION

Moraxella (Branhamella) catarrhalis bacteria are Gram-negative diplococcal pathogens which are carried asymptomatically in the healthy human respiratory tract. In recent years, *M. catarrhalis* has been recognized as an important causative agent of otitis media. In addition, *M. catarrhalis* has been associated with sinusitis, conjunctivitis, and urogenital infections, as well as with a number of inflammatory diseases of the lower respiratory tract in children and adults, including pneumonia, chronic bronchitis, tracheitis, and emphysema (refs. 1 to 8). (Throughout this application, various references are cited in parentheses to describe more fully the state of the art to which this invention pertains. Full bibliographic information for each citation is found at the end of the specification, immediately preceding the claims. The disclosures of these references are hereby incorporated by reference

into the present disclosure). Occasionally, *M. catarrhalis* invades to cause septicaemia, arthritis, endocarditis, and meningitis (refs. 9 to 13).

Otitis media is one of the most common illnesses of early childhood and approximately 80% of all children suffer at least one middle ear infection before the age of three (ref. 14). Chronic otitis media has been associated with auditory and speech impairment in children, and in some cases, has been associated with learning disabilities. Conventional treatments for otitis media include antibiotic administration and surgical procedures, including tonsillectomies, adenoidectomies, and tympanocentesis. In the United States, treatment costs for otitis media are estimated to be between one and two billion dollars per year.

In otitis media cases, *M. catarrhalis* commonly is co-isolated from middle ear fluid along with *Streptococcus pneumoniae* and non-typable *Haemophilus influenzae*, which are believed to be responsible for 50% and 30% of otitis media infections, respectively. *M. catarrhalis* is believed to be responsible for approximately 20% of otitis media infections (ref. 15). Epidemiological reports indicate that the number of cases of otitis media attributable to *M. catarrhalis* is increasing, along with the number of antibiotic-resistant isolates of *M. catarrhalis*. Thus, prior to 1970, no β-lactamase-producing *M. catarrhalis* isolates had been reported, but since the mid-seventies, an increasing number of β-lactamase-expressing isolates have been detected. Recent surveys suggest that 75% of clinical isolates produce β-lactamase (ref. 16, 26).

Iron is an essential nutrient for the growth of many bacteria. Several bacterial species, including *M. catarrhalis*, obtain iron from the host by using transferrin receptor proteins to capture transferrin. A number of bacteria including *Neisseria meningitidis*

(ref. 17), *N. gonorrhoeae* (ref. 18), *Haemophilus influenzae* (ref. 19), as well as *M. catarrhalis* (ref. 20), produce outer membrane proteins which specifically bind human transferrin. The expression of these proteins is regulated by the amount of iron in the environment.

The two transferrin receptor proteins of *M. catarrhalis*, designated transferrin binding protein 1 (Tbp1) and transferrin binding protein 2 (Tbp2), have molecular weights of 115 kDa (Tbp1) and approximately 80 to 90 kDa (Tbp2). Unlike the transferrin receptor proteins of other bacteria which have an affinity for apotransferrin, the *M. catarrhalis* Tbp2 receptors have a preferred affinity for iron-saturated (i.e., ferri-) transferrin (ref. 21).

M. catarrhalis infection may lead to serious disease. It would be advantageous to provide a recombinant source of transferrin binding proteins as antigens in immunogenic preparations including vaccines, carriers for other antigens and immunogens and the generation of diagnostic reagents. The genes encoding transferrin binding proteins and fragments thereof are particularly desirable and useful in the specific identification and diagnosis of *Moraxella* and for immunization against disease caused by *M. catarrhalis* and for the generation of diagnostic reagents.

SUMMARY OF THE INVENTION

The present invention is directed towards the provision of purified and isolated nucleic acid molecules encoding a transferrin receptor of a strain of *Moraxella* or a fragment or an analog of the transferrin receptor protein. The nucleic acid molecules provided herein are useful for the specific detection of strains of *Moraxella* and for diagnosis of infection by *Moraxella*. The purified and isolated nucleic acid

molecules provided herein, such as DNA, are also useful for expressing the *tbp* genes by recombinant DNA means for providing, in an economical manner, purified and isolated transferrin receptor proteins as well as subunits, fragments or analogs thereof. The transferrin receptor, subunits or fragments thereof or analogs thereof, as well as nucleic acid molecules encoding the same and vectors containing such nucleic acid molecules, are useful in immunogenic compositions for vaccinating against diseases caused by *Moraxella*, the diagnosis of infection by *Moraxella* and as tools for the generation of immunological reagents. Monoclonal antibodies or mono-specific antisera (antibodies) raised against the transferrin receptor protein, produced in accordance with aspects of the present invention, are useful for the diagnosis of infection by *Moraxella*, the specific detection of *Moraxella* (in, for example, *in vitro* and *in vivo* assays) and for the treatment of diseases caused by *Moraxella*.

In accordance with one aspect of the present invention, there is provided a purified and isolated nucleic acid molecule encoding a transferrin receptor protein of a strain of *Moraxella*, more particularly, a strain of *M. catarrhalis*, specifically *M. catarrhalis* strain 4223, Q8 or R1, or a fragment or an analog of the transferrin receptor protein.

In one preferred embodiment of the invention, the nucleic acid molecule may encode only the Tbp1 protein of the *Moraxella* strain or only the Tbp2 protein of the *Moraxella* strain. In another preferred embodiment of the invention, the nucleic acid may encode a fragment of the transferrin receptor protein of a strain of *Moraxella* having an amino acid sequence which is conserved.

In another aspect of the present invention, there is provided a purified and isolated nucleic acid

molecule having a DNA sequence selected from the group consisting of (a) a DNA sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, 8, 45 or 46) or the complementary DNA sequence thereto; (b) a DNA sequence encoding an amino acid sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 9, 10, 11, 12, 13, 14, 15, 16 or 47) or the complementary DNA sequence thereto; and (c) a DNA sequence which hybridizes under stringent conditions to any one of the DNA sequences defined in (a) or (b). The DNA sequence defined in (c) preferably has at least about 90% sequence identity with any one of the DNA sequences defined in (a) and (b). The DNA sequence defined in (c) may be that encoding the equivalent transferrin receptor protein from another strain of *Moraxella*.

In an additional aspect, the present invention includes a vector adapted for transformation of a host, comprising a nucleic acid molecule as provided herein and may have the characteristics of a nucleotide sequence contained within vectors LEM3-24, pLEM3, pLEM25, pLEM23, SLRD-A, DS-1698-1-1, DS-1754-1, pSLRD2, pSLRD3, pSLRD4 and pSLRD5.

The vector may be adapted for expression of the encoded transferrin receptor, fragments or analogs thereof, in a heterologous or homologous host, in either a lipidated or non-lipidated form. Accordingly, a further aspect of the present invention provides an expression vector adapted for transformation of a host comprising a nucleic acid molecule as provided herein and expression means operatively coupled to the nucleic acid molecule for expression by the host of the transferrin receptor protein or the fragment or analog of the transferrin receptor protein. In specific embodiments of this aspect of the invention, the nucleic acid molecule may encode substantially all the transferrin receptor protein, only the Tbp1 protein,

only the Tbp2 protein of the *Moraxella* strain or fragments of the Tbp1 or Tbp2 proteins. The expression means may include a promoter and a nucleic acid portion encoding a leader sequence for secretion from the host of the transferrin receptor protein or the fragment or the analog of the transferrin receptor protein. The expression means also may include a nucleic acid portion encoding a lipidation signal for expression from the host of a lipidated form of the transferrin receptor protein or the fragment or the analog of the transferrin receptor protein. The host may be selected from, for example, *Escherichia coli*, *Bordetella*, *Bacillus*, *Haemophilus*, *Moraxella*, fungi, yeast or baculovirus and Semliki Forest virus expression systems may be used. In a particular embodiment, the plasmid adapted for expression of Tbp1 is pLEM29 and that for expression of Tbp2 is pLEM33. Further vectors include pLEM-37, SLRD35-A and SLRD-35-B.

In an additional aspect of the invention, there is provided a transformed host containing an expression vector as provided herein. The invention further includes a recombinant transferrin receptor protein or fragment or analog thereof of a strain of *Moraxella* producible by the transformed host.

Such recombinant transferrin receptor protein may be provided in substantially pure form according to a further aspect of the invention, which provides a method of forming a substantially pure recombinant transferrin receptor protein, which comprises growing the transformed host provided herein to express a transferrin receptor protein as inclusion bodies, purifying the inclusion bodies free from cellular material and soluble proteins, solubilizing transferrin receptor protein from the purified inclusion bodies, and purifying the transferrin receptor protein free from other solubilized materials. The substantially pure

recombinant transferrin receptor protein may comprise Tbp1 alone, Tbp2 alone or a mixture thereof. The recombinant protein is generally at least about 70% pure, preferably at least about 90% pure.

Further aspects of the present invention, therefore, provide recombinantly-produced Tbp1 protein of a strain of *Moraxella* devoid of the Tbp2 protein of the *Moraxella* strain and any other protein of the *Moraxella* strain and recombinantly-produced Tbp2 protein of a strain of *Moraxella* devoid of the Tbp1 protein of the *Moraxella* strain and any other protein of the *Moraxella* strain. The *Moraxella* strain may be *M. catarrhalis* 4223 strain, *M. catarrhalis* Q8 strain or *M. catarrhalis* R1 strain.

In accordance with another aspect of the invention, an immunogenic composition is provided which comprises at least one active component selected from at least one nucleic acid molecule as provided herein and at least one recombinant protein as provided herein, and a pharmaceutically acceptable carrier therefor or vector therefor. The at least one active component produces an immune response when administered to a host.

The immunogenic compositions provided herein may be formulated as vaccines for *in vivo* administration to a host. For such purpose, the compositions may be formulated as a microparticle, capsule, ISCOM or liposome preparation. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces. The immunogenic compositions of the invention (including vaccines) may further comprise at least one other immunogenic or immunostimulating material and the immunostimulating material may be at least one adjuvant or at least one cytokine. Suitable adjuvants for use in the present invention include (but are not limited to) aluminum phosphate, aluminum

hydroxide, QS21, Quil A, derivatives and components thereof, ISCOM matrix, calcium phosphate, calcium hydroxide, zinc hydroxide, a glycolipid analog, an octadecyl ester of an amino acid, a muramyl dipeptide polyphosphazene, ISCOPREP, DC-chol, DDBA and a lipoprotein. Advantageous combinations of adjuvants are described in copending United States Patent Applications Nos. 08/261,194 filed June 16, 1994 and 08/483,856, filed June 7, 1995, assigned to the assignee hereof and the disclosures of which are incorporated herein by reference thereto (WO 95/34308).

In accordance with another aspect of the invention, there is provided a method for generating an immune response in a host, comprising the step of administering to a susceptible host, such as a human, an effective amount of the immunogenic composition provided herein. The immune response may be a humoral or a cell-mediated immune response and may provide protection against disease caused by *Moraxella*. Hosts in which protection against disease may be conferred include primates, including humans.

In a further aspect, there is provided a live vector for delivery of transferrin receptor to a host, comprising a vector containing the nucleic acid molecule as described above. The vector may be selected from *Salmonella*, BCG, adenovirus, poxvirus, vaccinia and poliovirus.

The nucleic acid molecules provided herein are useful in diagnostic applications. Accordingly, in a further aspect of the invention, there is provided a method of determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising the steps of:

(a) contacting the sample with a nucleic acid molecule as provided herein to produce duplexes comprising the nucleic acid molecule and any nucleic

acid molecule encoding the transferrin receptor protein of a strain of *Moraxella* present in the sample and specifically hybridizable therewith; and

(b) determining the production of the duplexes.

In addition, the present invention provides a diagnostic kit for determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising:

(a) a nucleic acid molecule as provided herein;

(b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the nucleic acid molecule and any such nucleic acid present in the sample and hybridizable with the nucleic acid molecule; and

(c) means for determining production of the duplexes.

The invention further includes the use of the nucleic acid molecules and proteins provided herein as medicines. The invention additionally includes the use of the nucleic acid molecules and proteins provided herein in the manufacture of medicaments for protection against infection by strains of *Moraxella*.

Advantages of the present invention include:

- an isolated and purified nucleic acid molecule encoding a transferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the transferrin receptor protein;
- recombinantly-produced transferrin receptor proteins, including Tbp1 and Tbp2, free from each other and other *Moraxella* proteins; and
- diagnostic kits and immunological reagents for specific identification of *Moraxella*.

BRIEF DESCRIPTION OF DRAWINGS

The present invention will be further understood from the following description with reference to the

drawings, in which:

Figure 1 shows the amino acid sequences (SEQ ID Nos: 17 and 18) of a conserved portion of Tbp1 proteins used for synthesis of degenerate primers used for PCR amplification of a portion of the *M. catarrhalis* 4223 *tbpA* gene;

Figure 2 shows a restriction map of clone LEM3-24 containing the *tbpA* and *tbpB* genes from *M. catarrhalis* isolate 4223;

Figure 3 shows a restriction map of the *tbpA* gene for *M. catarrhalis* 4223;

Figure 4 shows a restriction map of the *tbpB* gene for *M. catarrhalis* 4223;

Figure 5 shows the nucleotide sequence of the *tbpA* gene (SEQ ID No: 1 - entire sequence and SEQ ID No: 2 - coding sequence) and the deduced amino acid sequence of the Tbp1 protein from *M. catarrhalis* 4223 (SEQ ID No: 9 - full length and SEQ ID No: 10 - mature protein). The leader sequence (SEQ ID No: 19) is shown by underlining;

Figure 6 shows the nucleotide sequence of the *tbpB* gene (SEQ ID No: 3 - entire sequence and SEQ ID No: 4 - coding sequence) and the deduced amino acid sequence of the Tbp2 protein from *M. catarrhalis* 4223 (SEQ ID Nos: 11 - full length and SEQ ID No: 12 - mature protein). The leader sequence (SEQ ID No: 20) is shown by underlining;

Figure 7 shows a restriction map of clone SLRD-A containing the *tbpA* and *tbpB* genes from *M. catarrhalis* Q8;

Figure 8 shows a restriction map of the *tbpA* gene from *M. catarrhalis* Q8;

Figure 9 shows a restriction map of the *tbp*B gene from *M. catarrhalis* Q8;

Figure 10 shows the nucleotide sequence of the *tbpA* gene (SEQ. ID No: 5 - entire sequence and SEQ ID No: 6 - coding sequence) and the deduced amino acid sequence of

the Tbp1 protein from *M. catarrhalis* Q8 (SEQ ID No: 13 - full length and SEQ ID No: 14 - mature protein);

Figure 11 shows the nucleotide sequence of the *tbpB* gene (SEQ. ID No: 7 - entire sequence and SEQ ID No: 8 - coding sequence) and the deduced amino acid sequence of the Tbp2 protein from *M. catarrhalis* Q8 (SEQ ID No: 15 - full length and SEQ ID No: 16 - mature protein);

Figure 12 shows a comparison of the amino acid sequences of Tbp1 from *M. catarrhalis* strain 4223 (SEQ ID No: 9) and Q8 (SEQ ID No: 13), *H. influenzae* strain Eagan (SEQ ID No: 21), *N. meningitidis* strains B16B6 (SEQ ID No: 22) and M982 (SEQ ID No: 23), and *N. gonorrhoeae* strain FA19 (SEQ ID No: 24). Dots indicate identical residues and dashes have been inserted for maximum alignment;

Figure 13 shows a comparison of the amino acid sequences of Tbp2 from *M. catarrhalis* isolate 4223 (SEQ ID No: 11) and Q8 (SEQ ID No: 15), *H. influenzae* strain Eagan (SEQ ID No: 25), *N. meningitidis* strains B16B6 (SEQ ID No: 26) and M918 (SEQ ID No: 27), and *N. gonorrhoeae* strain FA19 (SEQ ID No: 28). Dots indicate identical residues and dashes have been inserted for maximum alignment;

Figure 14 shows the construction of plasmid pLEM29 for expression of recombinant Tbp1 protein from *E. coli;*

Figure 15 shows an SDS-PAGE analysis of the expression of Tbp1 protein by *E. coli* cells transformed with plasmid pLEM29;

Figure 16 shows a flow chart for purification of recombinant Tbp1 protein;

Figure 17 shows an SDS-PAGE analysis of purified recombinant Tbp1 protein;

Figure 18 shows the construction of plasmid pLEM33 and pLEM37 for expression of TbpA gene from *M. catarrhalis* 4223 in *E. coli* without and with a leader sequence respectively;

Figure 19 shows an SDS-PAGE analysis of the expression of rTbp2 protein by *E. coli* cells transformed with plasmid pLEM37;

Figure 20 shows the construction of plasmid sLRD35B for expression of the *tbpB* gene from *M. catarrhalis* Q8 in *E. coli* without a leader sequence, and the construction of plasmid SLRD35A for expression of the *tbpB* gene from *M. catarrhalis* Q8 in *E. coli* with a leader sequence. Restriction site B = BamHI; Bg = Bgl II; H = Hind III; R = EcoRI;

Figure 21 shows SDS PAGE analysis of the expression of rTbp2 protein in *E. coli* cells, transformed with plasmids SLRD35A and SLRD35B;

Figure 22 shows a flow chart for purification of recombinant Tbp2 protein from *E. coli;*

Figure 23, which includes Panels A and B, shows an SDS-PAGE analysis of the purification of recombinant Tbp2 protein from *M. catarrhalis* strains 4223 (Panel A) and Q8 (Panel B) from expression in *E. coli;*

Figure 24 shows the binding of Tbp2 to human transferrin;

Figure 25, which includes Panels A, B and C, shows the antigenic conservation of Tbp2 protein amongst strains of *M. catarrhalis;*

Figure 26 shows a restriction map of the *tbpB* gene for *M. catarrhalis* R1;

Figure 27 shows the nucleotide sequence of the *tbpB* gene (SEQ ID No: 45 - entire sequence and SEQ ID No: 46 - coding sequence) and the deduced amino acid sequence of the Tbp2 protein of *M. catarrhalis* R1 (SEQ ID No: 47); and

Figure 28 shows a comparison of the amino acid sequences of Tbp2 for *M. catarrhalis* 4223 (SEQ ID No: 21), Q8 (SEQ ID No: 15) and R1 (SEQ ID No: 47). Dots indicate identical residues and dashes have been inserted for maximum alignment. The asterisks indicate

stop codons.

GENERAL DESCRIPTION OF THE INVENTION

Any *Moraxella* strain may be conveniently used to provide the purified and isolated nucleic acid, which may be in the form of DNA molecules, comprising at least a portion of the nucleic acid coding for a transferrin receptor as typified by embodiments of the present invention. Such strains are generally available from clinical sources and from bacterial culture collections, such as the American Type Culture Collection.

In this application, the terms "transferrin receptor" (TfR) and "transferrin binding proteins" (Tbp) are used to define a family of Tbp1 and/or Tbp2 proteins which includes those having variations in their amino acid sequences including those naturally occurring in various strains of, for example, *Moraxella*. The purified and isolated DNA molecules comprising at least a portion coding for transferrin receptor of the present invention also include those encoding functional analogs of transferrin receptor proteins Tbp1 and Tbp2 of *Moraxella*. In this application, a first protein is a "functional analog" of a second protein if the first protein is immunologically related to and/or has the same function as the second protein. The functional analog may be, for example, a fragment of the protein, or a substitution, addition or deletion mutant thereof.

Chromosomal DNA from *M. catarrhalis* 4223 was digested with *Sau*3A in order to generate fragments within a 15 to 23 kb size range, and cloned into the *Bam*HI site of the lambda vector EMBL3. The library was screened with anti-Tbp1 guinea pig antisera, and a positive clone LEM3-24, containing an insert approximately 13.2 kb in size was selected for further analysis. Lysate from *E. coli* LE392 infected with LEM3-24 was found to contain a protein approximately 115 kDa

in size, which reacted on Western blots with anti-Tbp1 antisera. A second protein, approximately 80 kDa in size, reacted with the anti-Tbp2 guinea pig antisera on Western blots.

In order to localize the *tbpA* gene on the 13.2 kb insert of LEM3-24, degenerate PCR primers were used to amplify a small region of the putative *tbpA* gene of *M. catarrhalis* 4223. The sequences of the degenerate oligonucleotide primers were based upon conserved amino acid sequences within the Tbp1 proteins of several *Neisseria* and *Haemophilus* species and are shown in Figure 1 (SEQ ID Nos: 17 and 18). A 300 base-pair amplified product was generated and its location within the 4223 *tbpA* gene is indicated by bold letters in Figure 5 (SEQ ID No: 29). The amplified product was subcloned into the vector pCRII, labelled, and used to probe a Southern blot containing restriction-endonuclease digested clone LEM3-24 DNA. The probe hybridized to a 3.8 kb *Hin*dIII-*Hin*dIII, a 2.0 kb *Avr*II-*Avr*II, and 4.2 kb *Sal*I-*Sph*I fragments (Figure 2).

The 3.8 kb *Hin*dIII-*Hin*dIII fragment was subcloned into pACYC177, and sequenced. A large open reading frame was identified, and subsequently found to contain approximately 2 kb of the putative *tbpA* gene. The remaining 1 kb of the *tbpA* gene was obtained by subcloning an adjacent downstream *Hin*dIII-*Hin*dIII fragment into vector pACYC177. The nucleotide sequence of the *tbpA* gene from *M. catarrhalis* 4223 (SEQ ID Nos: 1 and 2), and the deduced amino acid sequence (SEQ ID No: 9 - full length; SEQ ID No: 10 mature protein) are shown in Figure 5.

Chromosomal DNA from *M. catarrhalis* strain Q8 was digested with Sau3A I and 15-23 kb fragments were ligated with BamHI arms of EMBL3. A high titre library was generated in *E. coli* LE392 cells and was screened using oligonucleotide probes based on the 4223 *tbpA*

sequence. Phage DNA was prepared and restriction enzyme analysis revealed that inserts of about 13-15 kb had been cloned. Phage clone SLRD-A was used to subclone fragments for sequence analysis. A cloning vector (pSKMA) was generated to facilitate cloning of the fragments and plasmids pSLRD1, pSLRD2, pSLRD3, pSLRD4 and pSLRD5 were generated which contain all of *tbpA* and most of *tbpB*. The nucleotide (SEQ ID Nos: 5 and 6) and deduced amino acid sequence (SEQ ID No: 13 - full length, SEQ ID No: 14 - mature protein) of the *tbpA* gene from strain Q8 are shown in Figure 10.

The deduced amino acid sequences for the Tbp1 protein encoded by the *tbpA* genes were found to share some nomology with the amino acid sequences encoded by genes from a number of *Neisseria* and *Haemophilus* species (Figure 12; SEQ ID Nos: 21, 22, 23 and 24).

Prior to the present discovery, *tbpA* genes identified in species of *Neisseria*, *Haemophilus*, and *Actinobacillus* have been found to be preceded by a *tbpB* gene with several conserved regions. The two genes typically are separated by a short intergenic sequence. However, a *tbpB* gene was not found upstream of the *tbpA* gene in *M. catarrhalis* 4223. In order to localize the *tbpB* gene within the 13.2 kb insert of clone LEM3-24, a denerate oligonucleotide probe was synthesized based upon an amino acid sequence EGGFYGP (SEQ ID No: 30), conserved among Tbp2 proteins of several species. The oligonucleotide was labelled and used to probe a Southern blot containing different restriction endonuclease fragments of clone LEM3-24. The probe hybridized to a 5.5 kb *Nhe*I-*Sal*I fragment, which subsequently was subcloned into pBR328, and sequenced. The fragment contained most of the putative *tbp*B gene, with the exception of the promoter region. The clone LEM3-24 was sequenced to obtain the remaining upstream sequence. The *tbp*B gene was located approximately 3 kb

downstream from the end of the *tbp*A gene, in contrast to the genetic organization of the *tbp*A and *tbp*B genes in *Haemophilus* and *Neisseria*. The nucleotide sequence (SEQ ID Nos: 3 and 4) of the *tbp*B gene from *M. catarrhalis* 4223 and the deduced amino acid sequence (SEQ ID Nos: 11, 12) are shown in Figure 6. The *tbpB* gene from *M. catarrhalis* Q8 was also cloned and sequenced. The nucleotide sequence (SEQ ID Nos: 7 and 8) and the deduced amino acid sequence (SEQ ID Nos: 15 and 16) are shown in Figure 11. The *tbpB* gene from *M. catarrhalis* R1 was also cloned and sequenced. The nucleotide sequence (SEQ ID Nos: 45 and 46) and the deduced amino acid sequence (SEQ ID No: 47) are shown in Figure 27. Regions of homology are evident between the *M. catarrhalis* Tbp2 amino acid sequences as shown in the comparative alignment of Figure 28 (SEQ ID Nos: 11, 15 and 47) and between the *M. catarrhalis* Tbp2 amino acid sequences and the Tbp2 sequences of a number of *Neisseria* and *Haemophilus* species, as shown in the comparative alignment in Figure 13 (SEQ ID Nos: 25, 26, 27, 28).

Cloned *tbpA* and *tbpB* genes were expressed in *E. coli* to produce recombinant Tbp1 and Tbp2 proteins free of other *Moraxella* proteins. These recombinant proteins were purified and used for immunization.

The antigenic conservation of Tbp2 protein amongst strains of *M. catarrhalis* was demonstrated by separation of the proteins in whole cell lysates of *M. catarrhalis* or strains of *E. coli* expressing recombinant Tbp2 proteins by SDS PAGE and antiserum immunoblotting with anti-4223 rTbp2 antiserum or anti-Q8 rTbp2 antiserum raised in guinea pigs. *M. catarrhalis* strains 3, 56, 135, 585, 4223, 5191, 8185 and ATCC 25240 were tested in this way and all showed specific reactivity with anti-4223 rTbp2 or anti-Q8 rTbp2 antibody (Figure 25).

In addition, the ability of anti-rTbp2 antibodies from one strain to recognize native or recombinant protein from the homologous or heterologous strain by ELISA is shown in Table 1 below.

Amino acid sequencing of the N-termini and cyanogen bromide fragments of transferrin receptor from *M. catarrhalis* 4223 was undertaken. Both N-termini of Tbp1 and Tbp2 were blocked. The putative signal sequences of Tbp1 and Tbp2 are indicated by underlining in Figures 5 and 6 (SEQ ID Nos: 19 and 20) respectively. The deduced amino acid sequences for the N-terminal region of Tbp2 suggests a lipoprotein structure.

Results shown in Tables 1 and 2 below illustrate the ability of anti-Tbp1 and anti-Tbp2 guinea pig antisera, produced by the immunization with Tbp1 or Tbp2, to lyze *M. catarrhalis*. The results show that the antisera produced by immunization with Tbp1 or Tbp2 protein isolated from *M. catarrhalis* isolate 4223 were bactericidal against a homologous non-clumping *M. catarrhalis* strain RH408 (a strain previously deposited in connection with United States Patent Application No. 08/328,589, assigned to the assignee hereof, (WO 96/12733) with the American Type Culture Collection, located at 1301 Parklawn Drive, Rockville, Maryland 20852, USA under the terms of the Budapest Treaty on December 13, 1994 under ATCC Deposit No. 55,637) derived from isolate 4223. In addition, antisera produced by immunization with Tbp1 protein isolated from *M. catarrhalis* 4223 were bactericidal against the heterologous non-clumping strain Q8 (a gift from Dr. M.G. Bergeron, Centre Hospitalier de l'Université Laval, St. Foy, Quebec). In addition, antiserum raised against recombinant Tbp2 (rTbp2) protein was bacteriacidal against the homologous strain of *M. catarrhalis*.

The ability of isolated and purified transferrin binding proteins to generate bactericidal antibodies is

in vivo evidence of utility of these proteins as vaccines to protect against disease caused by *Moraxella*.

Thus, in accordance with another aspect of the present invention, there is provided a vaccine against infection caused by *Moraxella* strains, comprising an immunogenically-effective amount of a transferrin binding protein from a strain of *Moraxella* and a physiologically-acceptable carrier therefor. Vaccine preparations may comprise antigenically or sequence divergent transferrin binding proteins.

The transferrin binding protein provided herein is useful as a diagnostic reagent, as an antigen for the generation of anti-transferrin protein binding antibodies, as an antigen for vaccination against the disease caused by species of *Moraxella* and for detecting infection by *Moraxella* and other such bacteria.

The transferrin binding protein provided herein may also be used as a carrier protein for haptens, polysaccharides or peptides to make conjugate vaccines against antigenic determinants unrelated to transferrin binding proteins. In additional embodiments of the present invention, therefore, the transferrin binding protein as provided herein may be used as a carrier molecule to prepare chimeric molecules and conjugate vaccines (including glycoconjugates) against pathogenic bacteria, including encapsulated bacteria. Thus, for example, glycoconjugates of the present invention may be used to confer protection against disease and infection caused by any bacteria having polysaccharide antigens including lipooligosaccharides (LOS) and PRP. Such bacterial pathogens may include, for example, *Haemophilus influenzae*, *Streptococcus pneumoniae*, *Escherichia coli*, *Neisseria meningitidis*, *Salmonella typhi*, *Streptococcus mutans*, *Cryptococcus neoformans*, *Klebsiella*, *Staphylococcus aureus* and *Pseudomonas aeruginosa*. Particular antigens which can be conjugated

to transferrin binding protein and methods to achieve such conjugations are described in U.S. Patent Application No. 08/433,522 filed November 23, 1993 (WO 94/12641), assigned to the assignee hereof and the disclosure of which is hereby incorporated by reference thereto.

In another embodiment, the carrier function of transferrin binding protein may be used, for example, to induce an immune response against abnormal polysaccharides of tumour cells, or to produce anti-tumour antibodies that can be conjugated to chemotherapeutic or bioactive agents.

The invention extends to transferrin binding proteins from *Moraxella catarrhalis* for use as an active ingredient in a vaccine against disease caused by infection with *Moraxella*. The invention also extends to a pharmaceutical vaccinal composition containing transferrin binding proteins from *Moraxella catarrhalis* and optionally, a pharmaceutically acceptable carrier and/or diluent.

In a further aspect the invention provides the use of transferrin binding proteins for the preparation of a pharmaceutical vaccinal composition for immunization against disease caused by infection with *Moraxella*.

It is clearly apparent to one skilled in the art, that the various embodiments of the present invention have many applications in the fields of vaccination, diagnosis, treatment of, for example, *Moraxella* infections and the generation of immunological and other diagnostic reagents. A further non-limiting discussion of such uses is further presented below.

1. Vaccine Preparation and Use

Immunogenic compositions, suitable to be used as vaccines, may be prepared from immunogenic transferrin receptor proteins, analogs and fragments thereof encoded by the nucleic acid molecules as well as the nucleic

acid molecules disclosed herein. The vaccine elicits an immune response which produces antibodies, including anti-transferrin receptor antibodies and antibodies that are opsonizing or bactericidal. Should the vaccinated subject be challenged by *Moraxella*, the antibodies bind to the transferrin receptor and thereby prevent access of the bacteria to an iron source which is required for viability. Furthermore, opsonizing or bactericidal anti-transferrin receptor antibodies may also provide protection by alternative mechanisms.

Immunogenic compositions, including vaccines, may be prepared as injectables, as liquid solutions or emulsions. The transferrin receptor proteins, analogs and fragments thereof and encoding nucleic acid molecules may be mixed with pharmaceutically acceptable excipients which are compatible with the transferrin receptor proteins, fragments, analogs or nucleic acid molecules. Such excipients may include water, saline, dextrose, glycerol, ethanol, and combinations thereof. The immunogenic compositions and vaccines may further contain auxiliary substances, such as wetting or emulsifying agents, pH buffering agents, or adjuvants, to enhance the effectiveness of the vaccines. Immunogenic compositions and vaccines may be administered parenterally, by injection subcutaneously, intradermally or intramuscularly. Alternatively, the immunogenic compositions provided according to the present invention, may be formulated and delivered in a manner to evoke an immune response at mucosal surfaces. Thus, the immunogenic composition may be administered to mucosal surfaces by, for example, the nasal or oral (intragastric) routes. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces. Some such targeting molecules include vitamin B12 and fragments of bacterial toxins, as

described in WO 92/17167 (Biotech Australia Pty. Ltd.), and monoclonal antibodies, as described in U.S. Patent No. 5,194,254 (Barber et al). Alternatively, other modes of administration, including suppositories and oral formulations, may be desirable. For suppositories, binders and carriers may include, for example, polyalkalene glycols or triglycerides. Oral formulations may include normally employed incipients such as, for example, pharmaceutical grades of saccharine, cellulose and magnesium carbonate. These compositions may take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain about 1 to 95% of the transferrin receptor proteins, fragments, analogs and/or nucleic acid molecules.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective, protective and immunogenic. The quantity to be administered depends on the subject to be treated, including, for example, the capacity of the individual's immune system to synthesize antibodies, and, if needed, to produce a cell-mediated immune response. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner. However, suitable dosage ranges are readily determinable by one skilled in the art and may be of the order of micrograms of the transferrin receptor proteins, analogs and fragments thereof and/or nucleic acid molecules. Suitable regimes for initial administration and booster doses are also variable, but may include an initial administration followed by subsequent administrations. The dosage of the vaccine may also depend on the route of administration and will vary according to the size of the host.

The nucleic acid molecules encoding the transferrin

receptor of *Moraxella* may be used directly for immunization by administration of the DNA directly, for example, by injection for genetic immunization or by constructing a live vector, such as *Salmonella*, BCG, adenovirus, poxvirus, vaccinia or poliovirus containing the nucleic acid molecules. A discussion of some live vectors that have been used to carry heterologous antigens to the immune system is contained in, for example, O'Hagan (ref 22). Processes for the direct injection of DNA into test subjects for genetic immunization are described in, for example, Ulmer et al. (ref. 23).

Immunogenicity can be significantly improved if the antigens are co-administered with adjuvants, commonly used as an 0.05 to 1.0 percent solution in phosphate - buffered saline. Adjuvants enhance the immunogenicity of an antigen but are not necessarily immunogenic themselves. Adjuvants may act by retaining the antigen locally near the site of administration to produce a depot effect facilitating a slow, sustained release of antigen to cells of the immune system. Adjuvants can also attract cells of the immune system to an antigen depot and stimulate such cells to elicit immune responses.

Immunostimulatory agents or adjuvants have been used for many years to improve the host immune responses to, for example, vaccines. Intrinsic adjuvants, such as lipopolysaccharides, normally are the components of killed or attenuated bacteria used as vaccines. Extrinsic adjuvants are immunomodulators which are typically non-covalently linked to antigens and are formulated to enhance the host immune responses. Thus, adjuvants have been identified that enhance the immune response to antigens delivered parenterally. Some of these adjuvants are toxic, however, and can cause undesirable side-effects, making them unsuitable for use

in humans and many animals. Indeed, only aluminum hydroxide and aluminum phosphate (collectively commonly referred to as alum) are routinely used as adjuvants in human and veterinary vaccines. The efficacy of alum in increasing antibody responses to diphtheria and tetanus toxoids is well established and an HBsAg vaccine has been adjuvanted with alum. While the usefulness of alum is well established for some applications, it has limitations. For example, alum is ineffective for influenza vaccination and inconsistently elicits a cell mediated immune response. The antibodies elicited by alum-adjuvanted antigens are mainly of the IgG1 isotype in the mouse, which may not be optimal for protection by some vaccinal agents.

A wide range of extrinsic adjuvants can provoke potent immune responses to antigens. These include saponins complexed to membrane protein antigens (immune stimulating complexes), pluronic polymers with mineral oil, killed mycobacteria and mineral oil, Freund's complete adjuvant, bacterial products, such as muramyl dipeptide (MDP) and lipopolysaccharide (LPS), as well as lipid A, and liposomes.

To efficiently induce humoral immune responses (HIR) and cell-mediated immunity (CMI), immunogens are often emulsified in adjuvants. Many adjuvants are toxic, inducing granulomas, acute and chronic inflammations (Freund's complete adjuvant, FCA), cytolysis (saponins and pluronic polymers) and pyrogenicity, arthritis and anterior uveitis (LPS and MDP). Although FCA is an excellent adjuvant and widely used in research, it is not licensed for use in human or veterinary vaccines because of its toxicity.

Desirable characteristics of ideal adjuvants include:

(1) lack of toxicity;

(2) ability to stimulate a long-lasting immune

response;

(3) simplicity of manufacture and stability in long-term storage;

(4) ability to elicit both CMI and HIR to antigens administered by various routes, if required;

(5) synergy with other adjuvants;

(6) capability of selectively interacting with populations of antigen presenting cells (APC);

(7) ability to specifically elicit appropriate T_H1 or T_H2 cell-specific immune responses; and

(8) ability to selectively increase appropriate antibody isotype levels (for example, IgA) against antigens.

U.S. Patent No. 4,855,283 granted to Lockhoff et al on August 8, 1989, which is incorporated herein by reference thereto, teaches glycolipid analogues including N-glycosylamides, N-glycosylureas and N-glycosylcarbamates, each of which is substituted in the sugar residue by an amino acid, as immuno-modulators or adjuvants. Thus, Lockhoff et al. 1991 (ref. 24) reported that N-glycolipid analogs displaying structural similarities to the naturally-occurring glycolipids, such as glycophospholipids and glycoglycerolipids, are capable of eliciting strong immune responses in both herpes simplex virus vaccine and pseudorabies virus vaccine. Some glycolipids have been synthesized from long chain-alkylamines and fatty acids that are linked directly with the sugars through the anomeric carbon atom, to mimic the functions of the naturally occurring lipid residues.

U.S. Patent No. 4,258,029 granted to Moloney, assigned to the assignee hereof and incorporated herein by reference thereto, teaches that octadecyl tyrosine hydrochloride (OTH) functions as an adjuvant when complexed with tetanus toxoid and formalin inactivated type I, II and III poliomyelitis virus vaccine. Also,

Nixon-George et al. 1990,(ref. 25) reported that octadecyl esters of aromatic amino acids complexed with a recombinant hepatitis B surface antigen, enhanced the host immune responses against hepatitis B virus.

2. Immunoassays

The transferrin receptor proteins, analogs and/or fragments thereof of the present invention are useful as immunogens, as antigens in immunoassays including enzyme-linked immunosorbent assays (ELISA), RIAs and other non-enzyme linked antibody binding assays or procedures known in the art for the detection of anti-*Moraxella*, transferrin receptor protein antibodies. In ELISA assays, the transferrin receptor protein, analogs and/or fragments corresponding to portions of TfR protein, are immobilized onto a selected surface, for example, a surface capable of binding proteins or peptides such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed transferrin receptor, analogs and/or fragments, a non-specific protein such as a solution of bovine serum albumin (BSA) or casein that is known to be antigenically neutral with regard to the test sample may be bound to the selected surface. This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by non-specific bindings of antisera onto the surface.

The immobilizing surface is then contacted with a sample, such as clinical or biological materials, to be tested in a manner conducive to immune complex (antigen/antibody) formation. This procedure may include diluting the sample with diluents, such as BSA, bovine gamma globulin (BGG) and/or phosphate buffered saline (PBS)/Tween. The sample is then allowed to incubate for from about 2 to 4 hours, at temperatures such as of the order of about 25° to 37°C. Following

incubation, the sample-contacted surface is washed to remove non-immunocomplexed material. The washing procedure may include washing with a solution such as PBS/Tween or a borate buffer.

Following formation of specific immunocomplexes between the test sample and the bound transferrin receptor protein, analogs and/or fragments and subsequent washing, the occurrence, and even amount, of immunocomplex formation may be determined by subjecting the immunocomplex to a second antibody having specificity for the first antibody. If the test sample is of human origin, the second antibody is an antibody having specificity for human immunoglobulins and in general IgG. To provide detecting means, the second antibody may have an associated activity such as an enzymatic activity that will generate, for example, a color development upon incubating with an appropriate chromogenic substrate. Quantification may then achieved by measuring the degree of color generation using, for example, a spectrophotometer.

3. Use of Sequences as Hybridization Probes

The nucleotide sequences of the present invention, comprising the sequence of the transferrin receptor gene, now allow for the identification and cloning of the transferrin receptor genes from any species of *Moraxella*.

The nucleotide sequences comprising the sequence of the transferrin receptor genes of the present invention are useful for their ability to selectively form duplex molecules with complementary stretches of other TfR genes. Depending on the application, a variety of hybridization conditions may be employed to achieve varying degrees of selectivity of the probe toward the other TfR genes. For a high degree of selectivity, relatively stringent conditions are used to form the duplexes, such as low salt and/or high temperature

conditions, such as provided by 0.02 M to 0.15 M NaCl at temperatures of between about 50°C to 70°C. For some applications, less stringent hybridization conditions are required such as 0.15 M to 0.9 M salt, at temperatures ranging from between about 20°C to 55°C. Hybridization conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be a method of choice depending on the desired results. In general, convenient hybridization temperatures in the presence of 50% formamide are: 42°C for a probe which is 95 to 100% homologous to the target fragment, 37°C for 90 to 95% homology and 32°C for 85 to 90% homology.

In a clinical diagnostic embodiment, the nucleic acid sequences of the TfR genes of the present invention may be used in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin and digoxigenin-labelling, which are capable of providing a detectable signal. In some diagnostic embodiments, an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of a radioactive tag may be used. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with samples containing TfR gene sequences.

The nucleic acid sequences of TfR genes of the present invention are useful as hybridization probes in solution hybridizations and in embodiments employing solid-phase procedures. In embodiments involving solid-

phase procedures, the test DNA (or RNA) from samples, such as clinical samples, including exudates, body fluids (e. g., serum, amniotic fluid, middle ear effusion, sputum, bronchoalveolar lavage fluid) or even tissues, is adsorbed or otherwise affixed to a selected matrix or surface. The fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes comprising the nucleic acid sequences of the TfR genes or fragments thereof of the present invention under desired conditions. The selected conditions will depend on the particular circumstances based on the particular criteria required depending on, for example, the G+C contents, type of target nucleic acid, source of nucleic acid, size of hybridization probe etc. Following washing of the hybridization surface so as to remove non-specifically bound probe molecules, specific hybridization is detected, or even quantified, by means of the label. It is preferred to select nucleic acid sequence portions which are conserved among species of *Moraxella*. The selected probe may be at least 18bp and may be in the range of about 30 to 90 bp.

4. Expression of the Transferrin Receptor Genes

Plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell may be used for the expression of the transferrin receptor genes in expression systems. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* may be transformed using pBR322 which contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, must also contain, or be modified to contain, promoters which can be used by the host cell for

expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host can be used as a transforming vector in connection with these hosts. For example, the phage in lambda GEM™-11 may be utilized in making recombinant phage vectors which can be used to transform host cells, such as *E. coli* LE392.

Promoters commonly used in recombinant DNA construction include the β-lactamase (penicillinase) and lactose promoter systems and other microbial promoters, such as the T7 promoter system as described in U.S. Patent No. 4,952,496. Details concerning the nucleotide sequences of promoters are known, enabling a skilled worker to ligate them functionally with genes. The particular promoter used will generally be a matter of choice depending upon the desired results. Hosts that are appropriate for expression of the transferrin receptor genes, fragments, analogs or variants thereof, may include *E. coli, Bacillus* species, *Haemophilus*, fungi, yeast, *Moraxella*, *Bordetella*, or the baculovirus expression system may be used.

In accordance with this invention, it is preferred to make the transferrin receptor protein, fragment or analog thereof, by recombinant methods, particularly since the naturally occurring TfR protein as purified from a culture of a species of *Moraxella* may include trace amounts of toxic materials or other contaminants. This problem can be avoided by using recombinantly produced TfR protein in heterologous systems which can be isolated from the host in a manner to minimize contaminants in the purified material. Particularly desirable hosts for expression in this regard include Gram positive bacteria which do not have LPS and are, therefore, endotoxin free. Such hosts include species of *Bacillus* and may be particularly useful for the

production of non-pyrogenic transferrin receptor, fragments or analogs thereof. Furthermore, recombinant methods of production permit the manufacture of Tbp1 or Tbp2 or respective analogs or fragments thereof, separate from one another which is distinct from the normal combined proteins present in *Moraxella*.

Biological Deposits

Certain vectors that contain at least a portion coding for a transferrin receptor protein from strains of *Moraxella catarrhalis* strain 4223 and Q8 and a strain of *M. catarrhalis* RH408 that are described and referred to herein have been deposited with the American Type Culture Collection (ATCC) located at 12301 Parklawn Drive, Rockville, Maryland, USA, pursuant to the Budapest Treaty and prior to the filing of this application. Samples of the deposited vectors and bacterial strain will become available to the public and the restrictions imposed on access to the deposits will be removed upon grant of a patent based upon this United States patent application. In addition, the deposit will be replaced if viable samples cannot be dispensed by the Depository. The invention described and claimed herein is not to be limited in scope by the biological materials deposited, since the deposited embodiment is intended only as an illustration of the invention. Any equivalent or similar vectors or strains that encode similar or equivalent antigens as described in this application are within the scope of the invention.

Deposit Summary

DEPOSIT	ATCC DESIGNATION	DATE DEPOSITED
Phage LEM3-24	97,381	December 4, 1995
Phage SLRD-A	97,380	December 4, 1995
Plasmid pLEM29	97,461	March 8, 1996
Plasmid pSLRD35A	97,833	January 13, 1997
Plasmid pLEM37	97,834	January 13, 1997
Strain RH408	55,637	December 9, 1994

EXAMPLES

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific Examples. These Examples are described solely for purposes of illustration and are not intended to limit the scope of the invention. Changes in form and substitution of equivalents are contemplated as circumstances may suggest or render expedient. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitations.

Methods of molecular genetics, protein biochemistry and immunology used but not explicitly described in this disclosure and these Examples are amply reported in the scientific literature and are well within the ability of those skilled in the art.

Example 1

This Example illustrates the preparation and immunization of guinea pigs with Tbp1 and Tbp2 proteins from *M. catarrhalis*.

Tbp1 and Tbp2 proteins were obtained as follows:

Iron-starved crude total membrane preparations were diluted to 4 mg protein/ml in 50 mM Tris.HCl-1M NaCl, pH 8, in a total volume of 384 ml. Membranes were solubilized by the addition of 8 ml each of 0.5M EDTA and 30% sarkosyl and samples were incubated for 2 hours at room temperature, with gentle agitation. Solubilized membranes were centrifuged at 10K rpm for 20 min. 15 ml of apo-hTf-Sepharose 4B were added to the supernatant, and incubated for 2 hours at room temperature, with gentle shaking. The mixture was added into a column. The column was washed with 50 ml of 50mM Tris.HCl-1 M NaCl-250mM guanidine hydrochloride, to remove contaminating proteins. Tbp2 was eluted from the column by the addition of 100 ml of 1.5M guanidine hydrochloride. Tbp1 was eluted by the addition of 100 ml of 3M guanidine hydrochloride. The first 20 ml fractions were dialyzed against 3 changes of 50 mM Tris.HCl, pH 8.0. Samples were stored at -20^{0}C, or dialyzed against ammonium bicarbonate and lyophilized.

Guinea pigs (Charles River) were immunized intramuscularly on day +1 with a 10 µg dose of Tbp1 or Tbp2 emulsified in complete Freund's adjuvant. Animals were boosted on days +14 and +29 with the same dose of protein emulsified in incomplete Freund's adjuvant. Blood samples were taken on day +42, and sera were used for analysis of bactericidal antibody activity. In addition, all antisera were assessed by immunoblot analysis for reactivity with *M. catarrhalis* 4223 proteins.

The bactericidal antibody activity of guinea pig anti-*M. catarrhalis* 4223 Tbp1 or Tbp2 antisera was determined as follows. A non-clumping *M. catarrhalis* strain RH408, derived from isolate 4223, was inoculated into 20 ml of BHI broth, and grown for 18 hr at 37°C, shaking at 170 rpm. One ml of this culture was used to

inoculate 20 ml of BHI supplemented with 25 mM ethylenediamine-di-hydroxyphenylacetic acid (EDDA; Sigma). The culture was grown to an OD_{578} of 0.5. The cells were diluted 1:200,000 in 140 mM NaCl, 93mM $NaHCO_3$, 2mM Na barbiturate, 4mM barbituric acid, 0.5mM $MgCl_2.6H_2O$, 0.4mM $CaCl_2.2H_2O$, pH 7.6 (Veronal buffer), containing 0.1% bovine serum albumin (VBS) and placed on ice. Guinea pig anti-*M. catarrhalis* 4223 Tbp1 or Tpb2 antisera, along with prebleed control antisera, were heated to 56°C for 30 min. to inactivate endogenous complement. Serial twofold dilutions of each antisera in VBS were added to the wells of a 96-well Nunclon microtitre plate (Nunc, Roskilde, Denmark). Dilutions started at 1:8, and were prepared to a final volume of 25 µL in each well. 25 µL of diluted bacterial cells were added to each of the wells. A guinea pig complement (Biowhittaker, Walkersville, MD) was diluted 1:10 in VBS, and 25 µL portions were added to each well. The plates were incubated at 37°C for 60 min, gently shaking at 70 rpm on a rotary platform. 50 µL of each reaction mixture were plated onto Mueller Hinton (Becton-Dickinson, Cockeysville, MD) agar plates. The plates were incubated at 37°C for 72 hr and the number of colonies per plate were counted. Bactericidal titres were assessed as the reciprocal of the highest dilution of antiserum capable of killing greater than 50% of bacteria compared with controls containing pre-immune sera. Results shown in Table 1 below illustrate the ability of the anti-Tbp1 and anti-Tbp2 guinea pig antisera to lyze *M. catarrhalis*.

Example 2

This Example illustrates the preparation of chromosomal DNA from *M. catarrhalis* strains 4223 and Q8.

M. catarrhalis isolate 4223 was inoculated into 100 ml of BHI broth, and incubated for 18 hr at 37°C with

shaking. The cells were harvested by centrifugation at 10,000 x g for 20 min. The pellet was used for extraction of *M. catarrhalis* 4223 chromosomal DNA.

The cell pellet was resuspended in 20 ml of 10 mM Tris-HCl (pH 7.5)-1.0 mM EDTA (TE). Pronase and SDS were added to final concentrations of 500 µg/ml and 1.0%, respectively, and the suspension was incubated at 37°C for 2 hr. After several sequential extractions with phenol, phenol:chloroform (1:1), and chloroform:isoamyl alcohol (24:1), the aqueous extract was dialysed, at 4°C, against 1.0 M NaCl for 4 hr, and against TE (pH 7.5) for a further 48 hr with three buffer changes. Two volumes of ethanol were added to the dialysate, and the DNA was spooled onto a glass rod. The DNA was allowed to air-dry, and was dissolved in 3.0 ml of water. Concentration was estimated, by UV spectrophotometry, to be about 290 µg/ml.

M. catarrhalis strain Q8 was grown in BHI broth as described in Example 1. Cells were pelleted from 50 ml of culture by centrifugation at 5000 rpm for 20 minutes, at 4°C. The cell pellet was resuspended in 10 ml of TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5) and proteinase K and SDS were added to final concentrations of 500 µg/ml and 1%, respectively. The sample was incubated at 37°C for 4 hours until a clear lysate was obtained. The lysate was extracted twice with Tris-saturated phenol/chloroform (1:1), and twice with chloroform. The final aqueous phase was dialysed for 24 hours against 2 X 1000 ml of 1 M NaCl at 4°C, changing the buffer once, and for 24 hours against 2 x 1000 ml of TE at 4°C, changing the buffer once. The final dialysate was precipitated with two volume of 100% ethanol. The DNA was spooled, dried and resuspended in 5 to 10 ml of TE buffer.

Example 3

This Example illustrates the construction of *M.*

catarrhalis chromosomal libraries in EMBL3.

A series of *Sau*3A restriction digests of chromosomal DNA, in final volumes of 10 µL each, were carried out in order to optimize the conditions necessary to generate maximal amounts of restriction fragments within a 15 to 23 kb size range. Using the optimized digestion conditions, a large-scale digestion was set up in a 100 µL volume, containing the following: 50 µL of chromosomal DNA (290 µg/ml), 33 µL water, 10 µL 10X *Sau*3A buffer (New England Biolabs), 1.0 µL BSA (10 mg/ml, New England Biolabs), and 6.3 µL *Sau*3A (0.04 U/µL). Following a 15 min. incubation at 37°C, the digestion was terminated by the addition of 10 µL of 100 mM Tris-HCl (pH 8.0)-10 mM EDTA-0.1% bromophenol blue-50% glycerol (loading buffer). Digested DNA was electrophoresed through a 0.5% agarose gel in 40 mM Tris acetate-2 mM $Na_2EDTA.2H_20$ (pH8.5)(TAE buffer) at 50 V for 6 hr. The region containing restriction fragments within a 15 to 23 kb molecular size range was excised from the gel, and placed into dialysis tubing containing 3.0 ml of TAE buffer. DNA was electroeluted from the gel fragment by applying a field strength of 1.0 V/cm for 18 hr. Electroeluted DNA was extracted once each with phenol and phenol:chloroform (1:1), and precipitated with ethanol. The dried DNA was dissolved in 5.0 µL water.

Size-fractionated chromosomal DNA was ligated with *Bam*HI-digested EMBL3 arms (Promega), using T4 DNA ligase in a final volume of 9 µL. The entire ligation mixture was packaged into lambda phage using a commercial packaging kit (Amersham), following manufacturer's instructions.

The packaged DNA library was amplified on solid media. 0.1 ml aliquots of *Escherichia coli* strain NM539 in 10 mM $MgSO_4$ (OD_{260} = 0.5) were incubated at 37°C for 15

min. with 15 to 25 µL of the packaged DNA library. Samples were mixed with 3 ml of 0.6% agarose containing 1.0% BBL trypticase peptone-0.5% NaCl (BBL top agarose), and mixtures were plated onto 1.5% agar plates containing 1.0% BBL trypticase peptone-0.5% NaCl, and incubated at 37°C for 18 hr. 3 ml quantities of 50 mM Tris-HCl (pH 7.5)-8 mM magnesium sulfate heptahydrate-100 mM NaCl-0.01% (w/v) gelatin (SM buffer) were added to each plate, and plates were left at 4°C for 7 hr. SM buffer containing phage was collected from the plates, pooled together, and stored in a screwcap tube at 4°C, with chloroform.

Chromosomal DNA from *M. catarrhalis* strain Q8 was digested with Sau3A I (0.1 unit/30 µg DNA) at 37°C for 30 minutes and size-fractionated on a 0.6% low melting point agarose gel. DNA fragments of 15-23 kb were excised and the DNA was electroeluted for 25 minutes in dialysis tubing containing TAE (40 mM Tris acetate pH 8.5, 2 mM EDTA) at 150 V. The DNA was extracted once with phenol/chloroform (1:1), precipitated, and resuspended in water. The DNA was ligated overnight with EMBL3 BamH I arms (Promega) and the ligation mixture was packaged using the Lambda *in vitro* packaging kit (Stratagene) and plated onto *E. coli* LE392 cells. The library was titrated and stored at 4°C in the presence of 0.3% chloroform.

Example 4

This Example illustrates screening of the *M. catarrhalis* libraries.

Ten µL aliquots of phage stock from the EMBL3/4223 sample prepared in Example 3 above were combined each with 100 µL of *E. coli* strain LE392 in 10 mM MgSO4 (OD_{260} = 0.5) (plating cells), and incubated at 37°C for 15 min. The samples were mixed with 3 ml each of BBL top agarose, and the mixtures were poured onto 1.5% agarose

plates containing 1% bacto tryptone-0.5% bacto yeast extract-0.05% NaCl (LB agarose; Difco) and supplemented with 200 μM EDDA. The plates were incubated at 37°C for 18 hr. Plaques were lifted onto nitrocellulose filters (Amersham Hybond-C Extra) using a standard protocol, and the filters were immersed into 5% bovine serum albumin (BSA; Boehringer) in 20 mM Tris-HCl (pH 7.5)-150 mM NaCl (TBS) for 30 min at room temperature, or 4°C overnight. Filters were incubated for at least 1 hr at room temperature, or 18 hr at 4°C, in TBS containing a 1/1000 dilution of guinea pig anti-*M. catarrhalis* 4223 Tbp1 antiserum. Following four sequential 10 min. washes in TBS with 0.05% Tween 20 (TBS-Tween), filters were incubated for 30 min. at room temperature in TBS-Tween containing a 1/4000 dilution of recombinant Protein G labelled with horseradish peroxidase (rProtein G-HRP; Zymed). Filters were washed as above, and submerged into CN/DAB substrate solution (Pierce). Color development was arrested by immersing the filters into water. Positive plaques were cored from the plates, and each placed into 0.5 ml of SM buffer containing a few drops of chloroform. The screening procedure was repeated two more times, until 100% of the lifted plaques were positive using the guinea pig anti-*M. catarrhalis* 4223 Tbp1 antiserum.

The EMBL3/Q8 library was plated onto LE392 cells on YT plates using 0.7% top agar in YT as overlay. Plaques were lifted onto nitrocellulose filters and the filters were probed with oligonucleotide probes labelled with $^{32}P\alpha$-dCTP (Random Primed DNA labeling kit, Boehringer Mannheim). The pre-hybridization was performed in sodium chloride/sodium citrate (SSC) buffer (ref. 27) at 37°C for 1 hour and the hybridization was performed at 42°C overnight. The probes were based upon an internal sequence of 4223 *tbpA*:

I R D L T R Y D P G

(Seq ID No. 31)

4236-RD 5' ATTCGAGACTTAACACGCTATGACCCTGGC 3'

(Seq ID No 32)

4237-RD 5' ATTCGTGATTTAACTCGCTATGACCCTGGT 3'

(Seq ID No 33).

Putative plaques were re-plated and submitted to second and third rounds of screening using the same procedures. Phage clone SLRD-A was used to subclone the *tfr* genes for sequence analysis.

Example 5

This Example illustrates immunoblot analysis of the phage lysates using anti-*M. catarrhalis* 4223 Tbp1 and Tbp2 antisera.

Proteins expressed by the phage eluants selected in Example 4 above were precipitated as follows. 60 µL of each phage eluant were combined with 200 µL *E. coli* LE392 plating cells, and incubated at 37°C for 15 min. The mixture was inoculated into 10 ml of 1.0% NZamine A-0.5% NaCl-0.1% casamino acids-0.5% yeast extract-0.2% magnesium sulfate heptahydrate (NZCYM broth), supplemented with 200 mM EDDA, and grown at 37°C for 18 hr, with shaking. DNAse was added to 1.0 ml of the culture, to a final concentration of 50 µg/ml, and the sample was incubated at 37°C for 30 min. Trichloroacetic acid was added to a final concentration of 12.5%, and the mixture was left on ice for 15 min. Proteins were pelleted by centrifugation at 13,000 x g for 10 min, and the pellet was washed with 1.0 ml of acetone. The pellet was air-dried and resuspended in 50 µL 4% SDS-20 mM Tris- HCl (pH 8.0)-0.2 mM EDTA (lysis buffer).

Following SDS-PAGE electrophoresis through an 11.5% gel, the proteins were transferred to Immobilon-P

filters (Millipore) at a constant voltage of 20 V for 18 hr, in 25mM Tris-HCl,220mM glycine-20% methanol (transfer buffer). Membranes were blocked in 5% BSA in TBS for 30 min. at room temperature. Blots were exposed either to guinea pig anti-*M. catarrhalis* 4223 Tbp1, or to guinea pig anti-*M. catarrhalis* 4223 Tbp2 antiserum, diluted 1/500 in TBS-Tween, for 2 hr at room temperature. Following three sequential 10 min. washes in TBS-Tween, membranes were incubated in TBS-Tween containing a 1/4000 dilution of rProtein G-HRP for 30 min. at room temperature. Membranes were washed as described above, and immersed into CN/DAB substrate solution. Color development was arrested by immersing blots into water.

Three EMBL3 phage clones expressed both a 115 kDa protein which reacted with anti-Tbp1 antiserum, and an 80 kDa protein, which reacted with anti-Tbp2 antiserum on Western blots and were thus concluded to contain genes encoding the transferrin receptor proteins of *Moraxella catarrhalis*.

Example 6

This Example illustrates the subcloning of the *M. catarrhalis* 4223 Tbp1 protein gene, *tbpA*.

Plate lysate cultures of the recombinant phage described in Example 5 were prepared by combining phage eluant and *E. coli* LE392 plating cells, to produce confluent lysis on LB agar plates. Phage DNA was extracted from the plate lysates using a Wizard Lambda Preps DNA Purification System (Promega), according to manufacturer's instructions.

The EMBL3 clone LM3-24 was found to contain a 13.2 kb insert, flanked by two *Sal*I sites. A probe to a *tbpA* gene was prepared and consisted of a 300 base pair amplified product generated by PCR using two degenerate oligonucleotide primers corresponding to an amino acid sequence of part of the Tbp1 protein (Figure 1). The

primer sequences were based upon the amino acid sequences NEVTGLG (SEQ ID No: 17) and GAINEIE (SEQ ID No: 18), which had been found to be conserved among the deduced amino acid sequences from several different *N. meningitidis* and *Haemophilus influenzae tbpA* genes. The amplified product was cloned into pCRII (Invitrogen, San Diego, CA) and sequenced. The deduced amino acid sequence shared homology with other putative amino acid sequences derived from *N. meningitidis* and *H. influenzae tbpA* genes (Figure 12). The subclone was linearized with *Not*I (New England Biolabs), and labelled using a digoxigenin random-labelling kit (Boehringer Mannheim), according to manufacturer's instructions. The concentration of the probe was estimated to be 2 ng/µL.

DNA from the phage clone was digested with *Hin*dIII, *Avr*II, *Sal*I/*Sph*I, or *Sal*I/*Avr*II, and electrophoresed through a 0.8% agarose gel. DNA was transferred to a nylon membrane (Genescreen Plus, Dupont) using an LKB VacuGene XL vacuum transfer apparatus (Pharmacia). Following transfer, the blot was air-dried, and pre-hybridized in 5X SSC-0.1% N-lauroylsarcosine-0.02% sodium dodecyl sulfate-1.0% blocking reagent (Boehringer Mannheim) in 10 mM maleic acid-15 mM NaCl (pH 7.5) (pre-hybridization solution). Labelled probe was added to the pre-hybridization solution to a final concentration of 6 ng/ml, and the blot was incubated in the probe solution at 42°C for 18 hr. The blot was washed twice in 2X SSC-0.1% SDS, for 5 min. each at room temperature, then twice in 0.1X SSC-0.1% SDS for 15 min. each at 60°C. Following the washes, the membrane was equilibrated in 100mM maleic acid-150 mM NaCl (pH 7.5) (buffer 1) for 1 min, then left in 1.0% blocking reagent (Boehringer Mannheim) in buffer 1 (buffer 2) for 60 min, at room temperature. The blot was exposed to anti-DIG-alkaline phosphatase (Boehringer Mannheim) diluted 1/5000 in buffer 2, for 30 min. at room temperature.

Following two 15 min. washes in buffer 1, the blot was equilibrated in 100 mM Tris-HCl (pH 9.5), 100 mM NaCl, 50 mM $MgCl_2$ (buffer 3) for 2 min. The blot was wetted with Lumigen PPD substrate (Boehringer-Mannheim), diluted 1/100 in buffer 3, then wrapped in Saran wrap, and exposed to X-ray film for 30 min. The probe hybridized to a 3.8 kb *Hin*dIII-*Hin*dIII, a 2.0 kb *Avr*II-*Avr*II, and a 4.2 kb *Sal*I-*Sph*I fragment.

In order to subclone the 3.8 kb *Hin*dIII-*Hind*III fragment into pACYC177, phage DNA from the EMBL3 clone, and plasmid DNA from the vector pACYC177 (New England Biolabs), were digested with *Hin*dIII, and fractionated by electrophoresis on a 0.8% agarose gel. The 3.8 kb *Hin*dIII-*Hin*dIII phage DNA fragment, and the 3.9 kb *Hin*dIII-*Hin*dIII pACYC177 fragment, were excised from the gel and purified using a Geneclean kit (Bio 101 Inc., LaJolla, CA), according to manufacturer's directions. Purified insert and vector were ligated together using T4 DNA ligase (New England Biolabs), and transformed into *E. coli* HB101 (Gibco BRL). A Qiagen Plasmid Midi-Kit (Qiagen) was used to extract and purify sequencing-quality DNA from one of the ampicillin-resistant/kanamycin-sensitive transformants, which was found to carry a 3.8 kb *Hin*dIII-*Hin*dIII insert. The subclone was named pLEM3. As described in Example 7, below, subsequent sequencing revealed that pLEM3 contained the first about 2.0 kb of *tbpA* sequence (Figures 2 and 5).

In order to subclone the remaining 1 kb of the *tbpA* gene, a 1.6 kb *Hin*dIII-*Hin*dIII fragment was subcloned into pACYC177 as described above, and transformed by electroporation into *E. coli* HB101 (Gibco BRL). A Midi-Plasmid DNA kit (Qiagen) was used to extract plasmid DNA from a putative kanamycin-sensitive transformant carrying a plasmid with a 1.6 kb *Hin*dIII-*Hin*dIII insert. The subclone was termed pLEM25. As described in

Example 7 below, sequencing revealed that pLEM25 contained the remaining 1 kb of the *tbpA* gene (Figure 2 and 5).

Example 7

This Example illustrates the subcloning of the *M. catarrhalis* 4223 *tbpB* gene.

As described above, in all *Neisseriae* and *Haemophilus* species examined prior to the present invention, *tbpB* genes have been found immediately upstream of the *tbpA* genes which share homology with the *tbpA* gene of *M. catarrhalis* 4223. However, the sequence upstream of *M. catarrhalis* 4223 did not correspond with other sequences encoding *tbpB*.

In order to localize the *tbpB* gene within the EMBL3 phage clone, a Southern blot was carried out using a degenerate probe from a highly conserved amino acid region within the Tbp2 protein. A degenerate oligonucleotide probe, was designed corresponding to the sequence encoding EGGFYGP (SEQ ID No: 30), which is conserved within the Tbp2 protein in a variety of *Neisseriae* and *Haemophilus* species. The probe was labelled with digoxigenin using an oligonucleotide tailing kit (Boehringer Mannheim), following the manufacturer's instructions. *Hin*dIII - digested EMBL3 clone DNA was fractionated through a 0.8% agarose gel, and transferred to a Geneclean Plus nylon membrane as described in Example 6. Following hybridization as described above, the membrane was washed twice in 2X SSC-0.1% SDS, for 5 min. each at room temperature, then twice in 0.1X SSC-0.1% SDS for 15 min. each, at 50°C. Detection of the labelled probe was carried out as described above. The probe hybridized to a 5.5 kb *Nhe*I-*Sal*I fragment.

The 5.5 kb *Nhe*I-*Sal*I fragment was subcloned into pBR328 as follows. LEM3-24 DNA, and pBR328 DNA, were digested with *Nhe*I-*Sal*I, and electrophoresed through

0.8% agarose. The 5.5 kb *Nhe*I-*Sal*I fragment, and the 4.9 kb pBR328 *Nhe*I-*Sal*I fragments were excised from the gel, and purified using a Geneclean kit as described in Example 6. The fragments were ligated together using T4 DNA ligase, and transformed into *E. coli* DH5. A Midi-Plasmid DNA kit (Qiagen) was used to extract DNA from an ampicillin resistant / tetracycline sensitive clone containing a 5.5 kb *Nhe*I-*Sal*I insert. This subclone was termed pLEM23. Sequencing revealed that pLEM23 contained 2 kb of the *tbpB* gene from *M. catarrhalis* 4223 (Figure 2).

Example 8:

This Example illustrates the subcloning of *M. catarrhalis* Q8 *tfr* genes.

The *M. catarrhalis* Q8 *tfr* genes were subcloned as follows. Phage DNA was prepared from plates. Briefly, the top agarose layer from three confluent plates was scraped into 9 ml of SM buffer (0.1 M NaCl, 0.2% $MgSO_4$, 50 mM Tris-HCl, pH 7.6, 0.01% gelatin) and 100 µl of chloroform was added. The mixture was vortexed for 10 sec, then incubated at room temperature for 2h. The cell debris was removed by centrifugation at 8000 rpm for 15 min at 4°C in an SS34 rotor (Sorvall model RC5C). The phage was pelleted by centrifugation at 35,000 rpm in a 70.1 Ti rotor at 10°C for 2h (Beckman model L8-80) and was resuspended in 500 µl of SM buffer. The sample was incubated at 4°C overnight, then RNAse and DNAse were added to final concentrations of 40 µg/ml and 10 µg/ml, respectively and the mixture incubated at 37°C for 1h. To the mixture were added 10 µl of 0.5 M EDTA and 5 µl of 10% SDS and the sample was incubated at 6°C for 15 min. The mixture was extracted twice with phenol/chloroform (1:1) and twice with chloroform and the DNA was precipitated by the addition of 2.5 volumes of absolute ethanol.

A partial restriction map was generated and fragments were subcloned using the external Sal I sites from EMBL3 and internal AvrII or EcoR I sites as indicated in Figure 4. In order to facilitate the subcloning, plasmid pSKMA was constructed which introduces a novel multiple cloning site into pBluescript.SK (Stratagene). Oligonucleotides were used to introduce restriction sites for Mst II, Sfi I, and Avr II between the Sal I and Hind III sites of pBluescript.SK:

```
                                         Sfi I
                  Sal I      Cla I    Mst II    Avr II HindIII
                    ↓          ↓        ↓    ↓    ↓       ↓
4639-RD   5' TCGACGGTAT CGATGGCC TTAG GGGC CTAGGA 3'
(SEQ ID No: 34)
4640-RD   3'      GCCATA GCTACCGG AATC CCCG GATCCTTCGA
(SEQ ID No: 35)
```

Plasmid pSLRD1 contains a ~1.5 kb Sal I-Avr II fragment cloned into pSKMA; plasmids pSLRD2 and pSLRD4 contain ~2 kb and 4 kb AvrII-AvrII fragments cloned into pSKMA, respectively and contain the complete *tbpA* gene. Plasmid pSLRD3 contains a ~2.3 kb AvrII-EcoR I fragment cloned into pSKMA and plasmid SLRD5 is a 22.7 kb EcoRI - EcoRI fragment cloned into pSKMA. These two clones contain the complete *tbpB* gene (Figure 7).

Example 9

This Example illustrates sequencing of the *M. catarrhalis tbp* genes.

Both strands of the *tbp* genes subcloned according to Examples 6 to 8 were sequenced using an Applied Biosystems DNA sequencer. The sequences of the *M. catarrhalis* 4223 and Q8 *tbpA* genes are shown in Figures 5 and 10 respectively. A derived amino acid sequence was compared with other Tbp1 amino acid sequences, including

those of *Neisseriae meningitidis*, *Neisseriae gonorrhoeae*, and *Haemophilus influenzae* (Figure 12). The sequence of the *M. catarrhalis* 4223 and Q8 *tbpB* genes are shown in Figures 6 and 11 respectively. In order to obtain sequence from the putative beginning of the *tbpB* gene of *M. catarrhalis* 4223, sequence data were obtained directly from the clone LEM3-24 DNA. This sequence was verified by screening clone DS-1754-1. The sequence of the translated *tbpB* genes from *M. catarrhalis* 4223 and Q8 shared homology with deduced Tbp2 amino acid sequences of *Neisseria meningitidis*, *Neisseria gonorrhoeae*, and *Haemophilus influenzae* (Figure 13).

Example 10

This Example illustrates the generation of an expression vector to produce recombinant Tbp1 protein. The construction scheme is shown in Figure 14.

Plasmid DNA from subclone pLEM3, prepared as described in Example 6, was digested with *Hin*dIII and *Bgl*I to generate a 1.84 kb *Bgl*I-*Hin*dIII fragment, containing approximately two-thirds of the *tbpA* gene. *Bam*HI was added to the digest to eliminate a comigrating 1.89kb *Bgl*I-*Hin*dIII vector fragment. In addition, plasmid DNA from the vector pT7-7 was digested with *Nde*I and *Hin*dIII. To create the beginning of the *tbpA* gene, an oligonucleotide was synthesized based upon the first 61 bases of the *tbpA* gene to the *Bgl*I site; an *Nde*I site was incorporated into the 5' end. Purified insert, vector and oligonucleotide were ligated together using T4 ligase (New England Biolabs), and transformed into *E. coli* DH5α. DNA was purified from one of the 4.4 kb ampicillin-resistant transformants containing correct restriction sites (pLEM27).

Purified pLEM27 DNA was digested with *Hin*dIII, ligated to the 1.6 kb *Hin*dIII-*Hin*dIII insert fragment

of pLEM25 prepared as described in Example 6, and transformed into *E. coli* DH5α. DNA was purified from an ampicillin-resistant transformant containing the correct restriction sites (pLEM29), and was transformed by electroporation into BL21 (DE3) (Novagen; Madison, WI) to produce *E. coli* pLEM29B-1.

A single isolated transformed colony was used to inoculate 100 ml of YT broth containing 100μg/ml ampicillin, and the culture was grown at 37^0C overnight, shaking at 200 rpm. 200 μl of the overnight culture were inoculated into 10 ml of YT broth containing 100μg/ml ampicillin, and the culture was grown at 37^0C to an OD_{578} of 0.35. The culture was induced by the addition of 30 μl of 100 mM IPTG, and the culture was grown at 37^0C for an additional 3 hours. One ml of culture was removed at the time of induction (t=0), and at t=1 hr and t=3 hrs. One ml samples were pelleted by centrifugation, and resuspended in 4%SDS-20 mM Tris.Cl, pH 8-200 μM EDTA (lysis buffer). Samples were fractionated on an 11.5% SDS-PAGE gel, and transferred onto Immobilon filters (Amersham). Blots were developed using anti-Tbp1 (*M. catarrhalis* 4223) antiserum, diluted 1:1000, as the primary antibody, and rproteinG conjugated with horseradish peroxidase (Zymed) as the secondary antibody. A chemiluminescent substrate (Lumiglo; Kirkegaard and Perry Laboratories, Gaithersburg, MD) was used for detection. Induced recombinant proteins were visible on the Coomassie-stained gels (Fig 15). The anti-Tbp1 (4223) antiserum recognized the recombinant proteins on Western blots.

Example 11

This Example illustrates the extraction and purification of recombinant Tbp1 of *M. catarrhalis* 4223.

Recombinant Tbp1 protein, which is contained in inclusion bodies, was purified from *E. coli* cells

expressing the *tbpA* gene (Example 10), by a procedure as shown in Figure 16. *E. coli* cells from a 500 ml culture, prepared as described in Example 10, were resuspended in 50 ml of 50 mM Tris-HCl, pH 8.0 containing 0.1 M NaCl and 5 mM AEBSF (protease inhibitor), and disrupted by sonication (3 x 10 min. 70% duty circle). The extract was centrifuged at 20,000 x g for 30 min. and the resultant supernatant which contained > 85% of the soluble proteins from *E. coli* was discarded.

The remaining pellet (Figure 16, PPT_1) was further extracted in 50 ml of 50 mM Tris, pH 8.0 containing 0.5% Triton X-100 and 10 mM EDTA. After centrifugation at 20,000 x g for 30 min., the supernatant containing residual soluble proteins and the majority of the membrane proteins was discarded.

The remaining pellet (Figure 16, PPT_2) was further extracted in 50 ml of 50 mM Tris, pH 8.0 containing 2M urea and 5 mM dithiothroitol (DTT). After centrifugation at 20,000 x g for 30 min., the resultant pellet (Figure 16, PPT_3) obtained after the above extraction contained the purified inclusion bodies.

The Tbp1 protein was solubilized from PPT3 in 50 mM Tris, pH 8.0, containing 6 M guanidine hydrochloride and 5 mM DTT. After centrifugation, the resultant supernatant was further purified on a Superdex 200 gel filtration column equilibrated in 50 mM Tris, pH 8.0, containing 2M guanidine hydrochloride and 5 mM DTT. The fractions were analyzed by SDS-PAGE and those containing purified Tbp1 were pooled. Triton X-100 was added to the pooled Tbpl fraction to a final concentration of 0.1%. The fraction was then dialyzed overnight at 4°C against 50 mM Tris, pH 8.0 and then centrifuged at 20,000 x g for 30 min. The protein remained soluble under these conditions and the purified Tbp1 was stored at -20° C. The purification procedure shown in Figure

16 produced Tbp1 protein that was at least 70% pure as determined by SDS-PAGE analysis (Figure 17).

Example 12

This Example illustrates the construction of an expression plasmid for rTbp2 of *M. catarrhalis* 4223 without a leader sequence.

The construction scheme for the plasmid expressing rTbp2 is shown in Figure 18. Oligonucleotides were used to construct the first approximately 58 bp of the *M. catarrhalis* 4223 *tbpB* gene encoding the mature protein. An *Nde*I site was incorporated into the 5' end of the oligonucleotides:

5'TATGTGTGGTGGCAGTGGTGGTTCAAATCCACCTGCTCCTACGCCCATT
CCAAATG (SEQ ID NO: 36) 3'

3'ACACACCACCGTCACCACCAAGTTTAGGTGGACGAGGATGCGGGTAAGG
TTTACGATC (SEQ ID NO: 37) 5'

An *Nhe*I-*Cla*I fragment, containing approximately 1kb of the *tbpB* gene from pLEM23, prepared as described in Example 7, was ligated to the above oligonucleotides and inserted into pT7-7 cut with *Nde*I-*Cla*I, generating pLEM31, which thus contains the 5′-half of *tbpB*. Oligonucleotides also were used to construct the last approximately 104 bp of the *tbpB* gene, from the *Ava*II site to the end of the gene. A *Bam*HI site was incorporated into the 3' end of the oligonucleotides:

5'GTCCAAATGCAAACGAGATGGGCGGGTCATTTACACACAACGCCGATG
ACAGCAAAGCCTCTGTGGTCTTTGGCACAAAAAGACAACAAGAAGTTAAGTAGTA
G (SEQ ID NO: 38) 3'

3'GTTTACGTTTGCTCTACCCGCCCAGTAAATGTGTGTTGCGGCTACTGTC
GTTTCGGAGACACCAGAAACCGTGTTTTTCTGTTGTTCTTCAATTCATCATCCTAG
(SEQ ID NO: 39) 5'

A *Cla*I-*Ava*II fragment from pLEM23, containing approximately 0.9 kb of the 3'-end of the *tbpB* gene, was ligated to the *Ava*II-*Bam*HI oligonucleotides, and inserted into pT7-7 cut with *Cla*I-*Bam*HI, generating pLEM32. The 1.0 kb *Nde*I-*Cla*I insert from pLEM31 and the 1.0 kb *Cla*I-*Bam*HI insert from pLEM32 were then inserted into pT7-7 cut with *Nde*I-*Bam*HI, generating pLEM33 which has a full-length *tbpB* gene under the direction of the T7 promoter.

DNA was purified from pLEM33 and transformed by electroporation into electrocompetent BL21(DE3) cells (Novagen; Madison, WI), to generate strain pLEM33B-1. Strain pLEM33B-1 was grown, and induced using IPTG, as described above in Example 10. Expressed proteins were resolved by SDS-PAGE and transferred to membranes suitable for immunoblotting. Blots were developed using anti-4223 Tbp2 antiserum, diluted 1:4000, as the primary antibody, and rprotein G conjugated with horseradish peroxidase (Zymed) as the secondary antibody. A chemiluminescent substrate (Lumiglo; Kirkegaard and Perry Laboratories, Gaithersburg, MD) was used for detection. Induced recombinant proteins were visible on the Coomassie blue-stained gels (Fig. 19). The anti-4223 Tbp2 antiserum recognized the recombinant proteins on Western blots.

Example 13

This Example illustrates the generation of an expression plasmid for rTbp2 of *M. catarrhalis* 4223 with a leader sequence.

The construction scheme is shown in Figure 18. Oligonucleotides containing the natural leader sequence of the *M. catarrhalis* 4223 *tbpB* gene were used to construct the first approximately 115 bp of the *tbpB* gene to the *Nhe*I site. An *Nde*I site was incorporated into the 5' end of the oligonucleotides:

5'TATGAAACACATTCCTTTAACCACACTGTGTGTGGCAATCTCTGCCGTC TTATTAACCGCTTGTGGTGGCAGTGGTGGTTCAAATCCACCTGCTCCTACGCCCAT TCCAAATG (SEQ ID NO: 40) 3'

3'ACTTTGTGTAAGGAAATTGGTGTGACACACACCGTTAGAGACGGCAGAA TAATTGGCGAACACCACCGTCACCACCAAGTTTAGGTGGACGAGGATGCGGGTAAG GTTTACGATC (SEQ ID NO: 41) 5'

The *Nde*I-*Nhe*I oligonucleotides were ligated to pLEM33 cut with *Nde*I-*Nhe*I, generating pLEM37, which thus contains a full-length 4223 *tbpB* gene encoding the Tbp2 protein with its leader sequence, driven by the T7 promoter.

DNA from pLEM37 was purified and transformed by electroporation into electrocompetent BL21(DE3) cells (Novagen; Madison, WI), to generate strain pLEM37B-2. pLEM37B-2 was grown, and induced using IPTG, as described above in Example 10. Expressed proteins were resolved by SDS-PAGE and transferred to membranes suitable for immunoblotting. Blots were developed using anti-4223 Tbp2 antiserum, diluted 1:4000, as the primary antibody, and rprotein G conjugated with horseradish peroxidase (Zymed) as the secondary antibody. A chemiluminescent substrate (Lumiglo; Kirkegaard and Perry Laboratories, Gaithersburg, MD) was used for detection. Induced recombinant proteins were visible on Coomassie-blue stained gels (Fig. 21). The anti-4223 Tbp2 antiserum recognized the recombinant proteins on Western blots.

Example 14

This Example illustrates the construction of an expression plasmid for rTbp2 of *M. catarrhalis* Q8 without a leader sequence.

The construction scheme for rTbp2 is shown in Figure 20. The 5'-end of the *tbpB* gene of *M. catarrhalis* Q8 was PCR amplified from the Cys^1 codon of

the mature protein through the Bsm I restriction site. An Nde I restriction site was introduced at the 5' end, for later cloning into pT7-7, and the final PCR fragment was 238 bp in length. The PCR primers are indicated below:

```
          NdeI   C   G   G   S   S   G   G   F   N
5' GAATTCCATATG TGT GGT GGG AGC TCT GGT GGT TTC AAT C
     3'  5247.RD (SEQ ID No: 42)
```

5' CCCATGGCAGGTTCTTGAATGCCTGAAACT 3' 5236.RD (SEQ ID No: 43).

The Q8 *tbpB* gene was subcloned in two fragments contained on plasmids SLRD3 and SLRD5, prepared as described in Example 8. Plasmid SLRD3-5 was constructed to contain the full-length *tbpB* gene by digesting SLRD5 with EcoR I and Dra I, which releases the 3'-end of *tbpB*, and inserting this ~ 619 bp fragment into SLRD3 which had been digested with EcoR I and Sma I. The 1.85 kb Bsm I-BamH I fragment from SLRD 3-5 was ligated with the 238 bp PCR fragment and inserted into pT7-7 that had been digested with Nde I and BamH I, generating plasmid SLRD35B. This plasmid thus contains the full-length *tbpB* gene without its leader sequence, under the direction of the T7 promoter. DNA from SLRD35B was purified and transformed by electroporation into electrocompetent BL21(DE3) cells to generate strain SLRD35BD which was grown and induced using IPTG, as described above in Example 10. Expressed proteins were resolved by SDS-PAGE and the induced Tbp2 protein was clearly visible by Coomassie blue staining (Fig. 19).

Example 15

This Example illustrates the generation of an expression plasmid for rTbp2 of *M. catarrhalis* Q8 with

a leader sequence.

The construction scheme for the rTbp2 is shown in Figure 20. The 5'-end of the Q8 *tbpB* gene was PCR amplified from the ATG start codon to the Bsm I restiction site. An Nde I site was engineered at the 5'-end, to facilitate cloning into the pT7-7 expression vector, and the final PCR fragment was 295 bp. The PCR primers are indicated below:

```
        Nde I   K   H   I   P   L   T
5' GAATTCCATATG AAA CAC ATT CCT TTA ACC 3'   5235.RD
(SEQ ID No: 44)

5' CCCATGGCAGGTTCTTGAATGCCTGAAACT       3'   5236.RD
(SEQ ID No: 43).
```

SLRD3-5 (Example 14) was digested with Bsm I and BamH I, generating a 1.85 kb fragment, which was ligated with the 295bp PCR fragment and ligated into pT7-7 that had been digested with Nde I and BamH I. The resulting plasmid SLRD35A thus contains the full-length Q8 *tbpB* gene with its endogenous leader sequence under the control of the T7 promoter. DNA from SLRD35A was purified and transformed by electroporation into electrocompetent BL21(DE3) cells to generate strain SLRD35AD which was grown and induced using IPTG, as described above in Example 10. Expressed proteins were resolved by SDS-PAGE and the induced Tbp2 protein was clearly visible by Coomassie blue staining (Fig. 19).

Example 16

This Example illustrates the extraction and purification of rTbp2 of *M. catarrhalis* 4223 and Q8 from *E. coli*.

pLEM37B (4223) and SLRD35AD (Q8) transformants were grown to produce Tbp2 in inclusion bodies and then the Tbp2 was purified according to the scheme in Figure

22. *E. coli* cells from a 500 mL culture, were resuspended in 50 mL of 50 mM Tris-HCl, pH 8.0 containing 5 mM AEBSF (protease inhibitor), and disrupted by sonication (3 x 10 min, 70% duty circle). The extract was centrifuged at 20,000 x *g* for 30 min and the resultant supernatant which contained > 95% of the soluble proteins from *E. coli* was discarded.

The remaining pellet (PPT_1) was further extracted in 50 mL of 50 mM Tris, pH 8.0 containing 0.5% Triton X-100 and 10 mM EDTA. The mixture was stirred at 4°C for at least 2 hours and then centrifuged at 20,000 x *g* for 30 min and the supernatant containing residual soluble proteins and the majority of the membrane proteins was discarded.

The resultant pellet (PPT_2) obtained after the above extraction contained the inclusion bodies. The Tbp2 protein was solubilized in 50 mM Tris, pH 8.0, containing 6 M guanidine and 5 mM DTT. After centrifugation, the resultant supernatant was further purified on a Superdex 200 gel filtration column equilibrated in 50 mM Tris, pH 8.0, containing 2 M guanidine and 5 mM DTT. The fractions were analyzed by SDS-PAGE and those containing purified Tbp2 were pooled. Triton X-100 was added to the pooled Tbp2 fraction to a final concentration of 0.1%. The fraction was then dialyzed overnight at 4°C against PBS, and then centrifuged at 20,000 x *g* for 30 min. The protein remained soluble under these conditions and the purified Tbp2 was stored at -20°C. Figure 22 shows the SDS PAGE analysis of fractions of the purification process for rTbp2 from strain 4223 (Panel A) and strain Q8 (Panel B). The rTbp2 was at least 70% pure.

Groups of five BALB/c mice were injected three times subcutaneously (s.c.) on days 1, 29 and 43 with purified rTbp2 (0.3 mg to 10 mg) from *M. catarrhalis* strains 4223 and Q8 in the presence or absence of $AlPO_4$

(1.5 mg per dose). Blood samples were taken on days 14, 28, 42 and 56 for analysing the anti-rTbp2 antibody titers by EIAs.

Groups of two rabbits and two guinea pigs (Charles River, Quebec) were immunized intramuscularly (i.m.) on day 1 with a 5 mg dose of purified rTbp2 protein emulsified in complete Freund's adjuvant (CFA). Animals were boosted on days 14 and 29 with the same dose of protein emulsified in incomplete Freund's adjuvant (IFA). Blood samples were taken on day 42 for analysing anti-rTbp2 antibody titers and bactericidal activity. Table 2 below shows the bactericidal activity of antibodies raised to the recombinant tranferrin binding proteins rTbp1 (4223), rTbp2 (4223) and rTbp2 (Q8), prepared as described in these Examples, against *M. catarrhalis* strains 4223 and Q8.

Example 17

This Example illustrates the binding of Tbp2 to human transferrin *in vitro*.

Transferrin-binding activity of Tbp2 was assessed according to the procedures of Schryvers and Lee (ref. 28) with modifications. Briefly, purified rTbp2 was subjected to discontinuous electrophoresis through 12.5% SDS-PAGE gels. The proteins were electrophoretically transferred to PVDF membrane and incubated with horseradish peroxidase-conjugated human transferrin (HRP-human transferrin, 1:50 dilution) (Jackson ImmunoResearch Labs Inc., Mississauga, Ontario) at 4°C for overnight. LumiGLO substrate (Kirkegaard & Perry Laboratories, Inc., Gaithersburg, MD) was used for chemiluminescent detection of HRP activity according to the manufacturer's instructions. Both 4223 rTbp2 and Q8 rTbp2 bind to human transferrin under these conditions, as shown in Figure 24.

Example 18

This Example illustrates antigenic conservation of

Tbp2 amongst *M. catarrhalis* strains.

Whole cell lysates of *M. catarrhalis* strains and *E. coli* strains expressing recombinant Tbp2 proteins were separated by SDS-PAGE and electrophoretically transferred to PVDF membrane. Guinea pig anti-4223 rTbp2 or anti-Q8 rTbp2 antisera were used as first antibody and alkaline phosphatase conjugated goat anti-guinea pig antibody was used as second antibody to detect Tbp2. *M. catarrhalis* strains 3, 56, 135, 585, 4223, 5191, 8185 and ATCC 25240 were tested and all showed specific reactivity with anti-4223 rTbp2 or anti-Q8 rTbp2 antibody (Figure 25).

Table 3 illustrates the ability of anti-rTbp2 antibodies from one *M. catarrhalis* strain to recognize native or recombinant protein from a homologous or heterologous *M. catarrhalis* strain.

Example 19

This Example illustrates PCR amplification of the *tbpB* gene from *M. catarrhalis* strain R1 and characterization of the amplified R1 *tbpB* gene.

Chromosomal DNA from *M. catarrhalis* strain R1 was prepared using standard techniques. The design of the oligonucleotide sense primer was based on a region approximately 274 bases upstream of the *M. catarrhalis* 4223 *tbpB* gene, and the antisense primer was based upon a region approximately 11 bases downstream of the end of 4223 *tbpB*. The following primers were used:

sense primer (4940): 5′ GATATAAGCACGCCCTACTT 3′
(SEQ ID No: 48)
antisense primer (4967): 5′ CCCATCAGCCAAACAAACATTGTGT 3′
(SEQ ID No: 49).

Each reaction tube contained 10 mM Tris-HCl (pH 8.85), 25 mM KCl, 5 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 800 mM dNTPs, 1.0 mg each of primers 4940 and 4967, 10 ng of R1 DNA, and 2.5 U Pwo DNA polymerase (Boehringer

Mannheim) in a total volume of 100 µl. The thermocycler was programmed for 5 min at 95°C, followed by 25 cycles of 95°C for 30 sec, 50°C for 45 sec, and 72°C for 2 min, and a 10 min final elongation elongation at 72°C. The amplified product was purified using a Geneclean (BIO 101) according to the manufacturer's instructions, and sequenced.

A partial restriction map of *M. catarrhalis* strain R1 *tbpB* prepared as just described is shown in Figure 26. The nucleotide and deduced amino acid sequences of the PCR amplified R1 *tbpB* gene are shown in Figure 27. The R1 *tbpB* gene encodes a 714 amino acid protein of molecular weight 76.8 kDa. The leader sequence of the R1 Tbp2 protein is identical to that of the 4223 and Q8 Tbp2 proteins. When the deduced R1 Tbp2 sequence was aligned with the 4223 Tbp2 sequence, it was found to be 83% identical and 88% homologous (Fig. 28). The conserved LEGGFYG (SEQ ID No: 50) epitope was present, as found in Tbp2 from other *M. catarrhalis* strains as well as the *H. influenzae* and *N. meningitidis* Tbp2 proteins.

SUMMARY OF THE DISCLOSURE

In summary of this disclosure, the present invention provides purified and isolated DNA molecules containing transferrin receptor genes of *Moraxella catarrhalis*, the sequences of these transferrin receptor genes, and the derived amino acid sequences thereof. The genes and DNA sequences are useful for diagnosis, immunization, and the generation of diagnostic and immunological reagents. Immunogenic compositions, including vaccines, based upon expressed recombinant Tbp1 and/or Tbp2, portions thereof, or analogs thereof, can be prepared for prevention of diseases caused by *Moraxella*. Modifications are possible within the scope of this invention.

TABLE I

BACTERICIDAL ANTIBODY TITRES FOR *M. CATARRHALIS* ANTIGENS

ANTIGEN[1]	SOURCE OF ANTISERA [2]	BACTERICIDAL TITRE[3] RH408[4]		BACTERICIDAL TITRE Q8[5]	
		Pre-Immune	Post-Immune	Pre-Immune	Post-Immune
TBP1	GP	< 3.0	4.2-6.9	< 3.0	4.4.-6.2
TBP2	GP	< 3.0	12.0-13.6	< 3.0	< 3.0-4.0

1 antigens isolated from *M. catarrhalis* 4223

2 GP = guinea pig

3 bactericidal titres: expressed in $\log_2$ as the dilution of antiserum capable of killing 50% of cells

4 *M. catarrhalis* RH408 is a non-clumping derivative of *M. catarrhalis* 4223

5 *M. catarrhalis* Q8 is a clinical isolate which displays a non-clumping phenotype

TABLE 2

	Bactericidal titre - RH408		Bactericidal titre - Q8	
Antigen	pre-immune	post-immune	pre-immune	post-immune
rTbp1 (4223)	< 3.0	< 3.0	< 3.0	< 3.0
rTbp2 (4223)	< 3.0	10-15	< 3.0	< 3.0
rTbp2 (Q8)	NT	NT	< 3.0	5.5-7.5

Antibody titres are expressed in log_2 as the dilution of antiserum capable of killing 50% of cells

NT = not tested

TABLE 3

ELISA titres for anti-rTbp2 antibodies recognizing native or rTbp2 from strain 4223 or rTbp2 from strain Q8

Coated antigen	Anti-rTbp2 (4223) Antibody Titres		Anti-rTbp2 (Q8) Antibody Titres	
	Rabbit antisera	Guinea pig antisera	Rabbit antisera	Guinea pig antisera
Native Tbp2 (4223)	409,600 204,800	1,638,400 1,638,400	25,600 25,600	51,200 102,400
rTbp2 (4223)	409,600 409,600	1,638,400 1,638,400	102,400 102,400	204,800 204,800
rTbp2 (Q8)	409,600 102,400	1,638,400 1,638,400	1,638,400 409,600	1,638,400 1,638,400

REFERENCES

1. Brorson, J-E., A. Axelsson, and S.E. Holm. 1976. Studies on *Branhamella catarrhalis* (*Neisseria catarrhalis*) with special reference to maxillary sinusitis. Scan. J. Infect. Dis. 8:151-155.

2. Catlin, B.W., 1990. *Branhamella catarrhalis*: an organism gaining respect as a pathogen. Clin. Microbiol. Rev. 3: 293-320.

3. Hager, H., A. Verghese, S. Alvarez, and S.L. Berk. 1987. *Branhamella catarrhalis* respiratory infections. Rev. Infect. Dis. 9:1140-1149.

4. McLeod, D.T., F. Ahmad, M.J. Croughan, and M.A. Calder. 1986. Bronchopulmonary infection due to *M. catarrhalis*. Clinical features and therapeutic response. Drugs 31(Suppl.3):109-112.

5. Nicotra, B., M. Rivera, J.I. Luman, and R.J. Wallace. 1986. *Branhamella catarrhalis* as a lower respiratory tract pathogen in patients with chronic lung disease. Arch.Intern.Med. 146:890-893.

6. Ninane, G., J. Joly, and M. Kraytman. 1978. Bronchopulmonary infection due to *Branhamella catarrhalis* 11 cases assessed by transtracheal puncture. Br.Med.Jr. 1:276-278.

7. Srinivasan, G., M.J. Raff, W.C. Templeton, S.J. Givens, R.C. Graves, and J.C. Mel. 1981. *Branhamella catarrhalis* pneumonia. Report of two cases and review of the literature. Am.Rev. Respir. Dis. 123:553-555.

8. West, M., S.L. Berk, and J.K. Smith. 1982. *Branhamella catarrhalis* pneumonia. South.Med. J. 75:1021-1023.

9. Christensen, J.J., and B. Bruun. 1985. Bacteremia caused by a beta-lactamase producing strain of *Branhamella catarrhalis*. Acta.Pathol. Microbiol. Immunol. Scand. Sect.B 93:273-275.

10. Craig, D.B., and P.A. Wehrle. 1983. *Branhamella catarrhalis* septic arthritis. J. Rheumatol. 10:985-986.

11. Guthrie, R., K. Bakenhaster, R.Nelson, and R. Woskobnick. 1988. *Branhamella catarrhalis* sepsis: a case report and review of the literature. J.Infect.Dis. 158:907-908.

12. Hiroshi, S., E.J. Anaissie, N.Khardori, and G.P. Bodey. 1988. *Branhamella catarrhalis* septicemia in patients with leukemia. Cancer 61:2315-2317.

13. O'Neill, J.H., and P.W. Mathieson. 1987. Meningitis due to *Branhamella catarrhalis*. Aust. N.Z. J. Med. 17:241-242.

14. Murphy, T.F. 1989. The surface of *Branhamella catarrhalis*: a systematic approach to the surface antigens of an emerging pathogen. Pediatr. Infect. Dis. J. 8:S75-S77.

15. Van Hare, G.F., P.A. Shurin, C.D. Marchant, N.A. Cartelli, C.E.Johnson, D. Fulton, S. Carlin, and C.H. Kim. Acute otitis media caused by *Branhamella catarrhalis*: biology and therapy. Rev. Infect. Dis. 9:16-27.

16. Jorgensen, J.H., Doern, G.V., Maher, L.A., Howell, A.W., and Redding, J.S., 1990 Antimicrobial resistance among respiratory isolates of *Haemophilus influenza, Moraxella catarrhalis*, and *Streptococcus* pneumoniae in the United States. Antibicrob. Agents Chemother. 34: 2075-2080.

17. Schryvers, A.B. and Morris, L.J. 1988 Identification and Characterization of the transferrin receptor from *Neisseria meningitidis*. Mol. Microbiol. 2:281-288.

18. Lee, B.C., Schryvers, A.B. Specificity of the lactoferrin and transferrin receptors in *Neisseria gonorrhoeae*. Mol. Microbiol. 1988; 2-827-9.

19. Schryvers, A.B. Characterization of the human transferrin and lactoferrin receptors in *Haemophilus influenzae*. Mol. Microbiol. 1988; 2: 467-72.

20. Schryvers, A.B. and Lee, B.C. (1988) Comparative analysis of the transferrin and lactoferrin binding proteins in the family *Neisseriaceae*. Can. J. Microbiol. 35, 409-415.

21. Yu, R. and Schryvers, A.B., 1993. The interaction between human transferrin and transferrin binding protein 2 from *Moraxella (Branhamella) catarrhalis* differs from that of other human pathogens. Microbiol. Pathogenesis, 15:433-445.

22. O'Hagan, 1992. Clin. Pharmokinet. 22:1

23. Ulmer et al., 1993. Curr. Opinion Invest. Drugs 2: 983-989.

24. Lockhoff, O., 1991. glycolipds as immunomoclutators: Synthesis and properits. Chem. Int. Ed. Engl. 30: 1611-1620.

25. Nixon-George, 1990. J. Immunol. 14: 4798-4802.

26. Wallace, R.J. Jr., Nash, D.R., and Steingrube, V.A. 1990. Antibiotic susceptibilites and drug resistance in *Moraxella (Branhaemella) catarrhalis*. Am. J. Med. 88 (5A): 465-50S.

27. F.M. Ausubel et al., Short protocols in Molecular Biology, Greene Publishing Associates and John Wiley and Sons.

28. Schryvers, A.B., Lee, B.C. 1989. Comparative analysis of the transferrin and lactoferrin binding proteins in the family *Neisseriaceae*. Can. J. Microbiol. 35: 409-415.

CLAIMS

What we claim is:

1. A purified and isolated nucleic acid molecule encoding a transferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the transferrin receptor protein.

2. The nucleic acid molecule of claim 1 wherein the transferrin receptor protein is the transferrin receptor binding protein 1 (Tbp1) of the *Moraxella* strain.

3. The nucleic acid molecule of claim 2 wherein the transferrin receptor protein is the transferrin receptor binding protein 2 (Tbp2) of the *Moraxella* strain.

4. The nucleic acid molecule of claim 1 wherein the strain of *Moraxella* is a strain of *Moraxella catarrhalis*.

5. The nucleic acid molecule of claim 4 wherein the strain of *Moraxella catarrhalis* is *Moraxella catarrhalis* 4223, Q8 or R1.

6. A purified and isolated nucleic acid molecule having a DNA sequence selected from the group consisting of:

(a) a DNA sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, 8, 45 or 46) or the complementary DNA sequence thereto;

(b) a DNA sequence encoding an amino acid sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 9, 10, 11, 12, 13, 14, 15, 16 or 47) or the complementary DNA sequence thereto; and

(c) a DNA sequence which hybridizes under stringent conditions to any one of the DNA sequences defined in (a) or (b).

7. The nucleic acid molecule of claim 6, wherein the DNA sequence defined in (c) has at least about 90% sequence identity with any one of the DNA sequences

defined in (a) or (b).

8. The nucleic acid molecule of claim 6 wherein the DNA sequence defined in (c) is that encoding the equivalent transferrin receptor protein from another strain of *Moraxella*.

9. A vector adapted for transformation of a host comprising the nucleic acid molecule of claim 1 or 6.

10. The vector of claim 9 encoding a fragment of a transferrin receptor protein and having the characteristics of a plasmid selected from the group consisting of pLEM3, pLEM25, pLEM23, DS-1698-1-1, DS-1754-1, pSLRD2, pSLRD3, pSLRD4 and pSLRD5.

11. The vector of claim 9 further comprising expression means operatively coupled to the nucleic acid molecule for expression by the host of said transferrin receptor protein of a strain of *Moraxella* or the fragment or the analog of the transferrin receptor protein.

12. The vector of claim 11 having the characteristics of plasmid pLEM-29, pLEM-33, pLEM-37, SLRD35-A and SLRD35-B.

13. A transformed host containing an expression vector as claimed in claim 11.

14. A method of forming a substantially pure recombinant transferrin receptor protein of a strain of *Moraxella*, which comprises:

growing the transformed host of claim 13 to express a transferrin receptor protein as inclusion bodies,

purifying the inclusion bodies free from cellular material and soluble proteins,

solubilizing transferrin receptor protein from the purified inclusion bodies, and

purifying the transferrin receptor protein free

from other solubilized materials.

15. The method of claim 14 wherein said transferrin receptor protein comprises Tbp1 alone, Tbp2 alone or a mixture of Tbp1 and Tbp2.

16. The method of claim 15 wherein said transferrin receptor protein is at least about 70% pure.

17. The method of claim 16 wherein said transferrin receptor protein is at least about 90% pure.

18. A recombinant transferrin receptor protein or fragment or analog thereof producible by the transformed host of claim 12.

19. The protein of claim 18 which is transferrin receptor binding protein 1 (Tbp1) of the *Moraxella* strain devoid of other proteins of the *Moraxella* strain.

20. The protein of claim 18 which is transferrin receptor binding protein 2 (Tbp2) of the *Moraxella* strain devoid of other proteins of the *Moraxella* strain.

21. The protein of claim 18 wherein the strain of *Moraxella* is a strain of *Moraxella catarrhalis*.

22. An immunogenic composition, comprising at least one active component selected from the group consisting of:

(A) a purified and isolated nucleic acid molecule encoding a transferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the transferrin receptor protein;

(B) a purified and isolated nucleic acid molecule having a DNA sequence selected from the group consisting of:

(a) a DNA sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, 8, 45 or 46) or the complementary DNA sequence thereto;

(b) a DNA sequence encoding an amino acid sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 9, 10, 11, 12, 13, 14, 15, 16 or 47) or the complementary DNA sequence thereto; and

(c) a DNA sequence which hybridizes under stringent conditions to any one of the DNA sequences defined in (a) or (b); or

(C) a recombinant transferrin receptor protein or fragment or analog thereof producible by a transformed host containing an expression vector comprising a nucleic acid molecule as defined in (A) or (B) and expression means operatively coupled to the nucleic acid molecule for expression by the host of the recombinant transferrin receptor protein or fragment or analog thereof;

and a pharmaceutically acceptable carrier therefor, said at least one active component producing an immune response when administered to a host.

23. A method for generating an immune response in a host, comprising administering to the host an immunoeffective amount of the immunogenic composition of claim 22.

24. A method of determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising the steps of:

(a) contacting the sample with the nucleic acid molecule of claim 1 or 6 to produce duplexes comprising the nucleic acid molecule and any said nucleic acid molecule encoding the transferrin receptor protein of a strain of *Moraxella* present in the sample and specifically hybridizable therewith; and

(b) determining production of the duplexes.

25. A diagnostic kit for determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising:

(a) the nucleic acid molecule of claim 1 or 6;

(b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the nucleic acid molecule and any said nucleic acid present in the sample and hybridizable with the nucleic acid molecule; and

(c) means for determining production of the duplexes.

AMINO ACID SEQUENCES OF A CONSERVED PORTION OF Tbp1 PROTEIN FOR CONSTRUCTION OF DEGENERATE PRIMERS USED IN PCR AMPLIFICATION OF A PORTION OF THE *M. cattarhalis* 4223 *tbp*A GENE.

N E V T G L G	SEQ ID NO: 17
G A I N E I E	SEQ ID NO: 18

FIG.1

09/142628

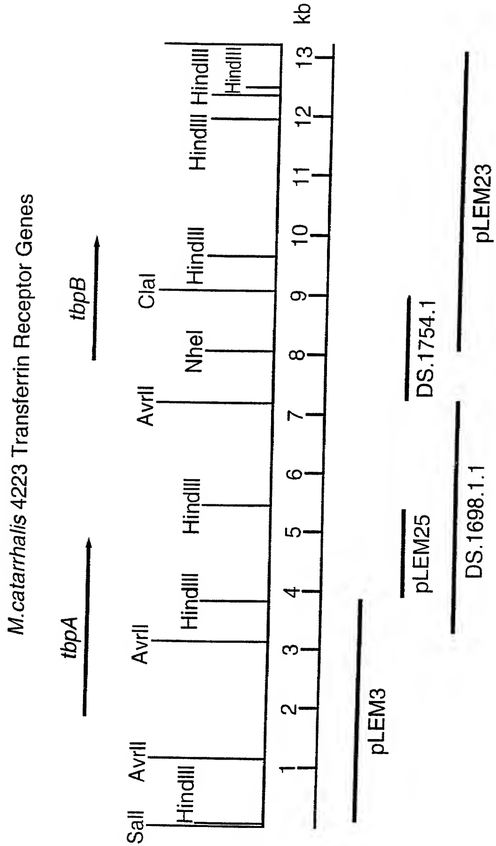


FIG.2

09/142028

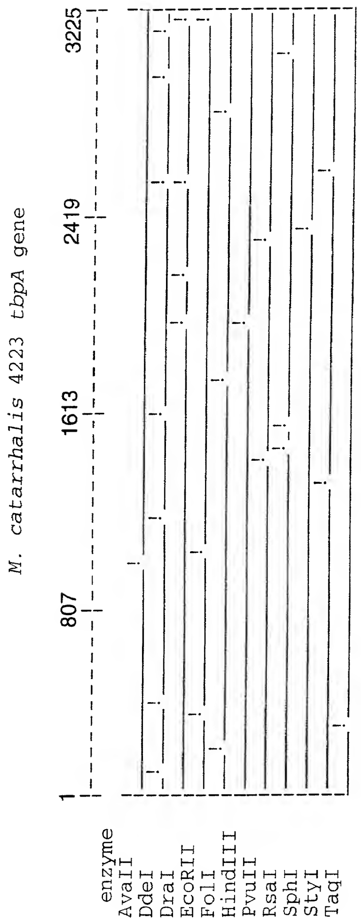


FIG.3

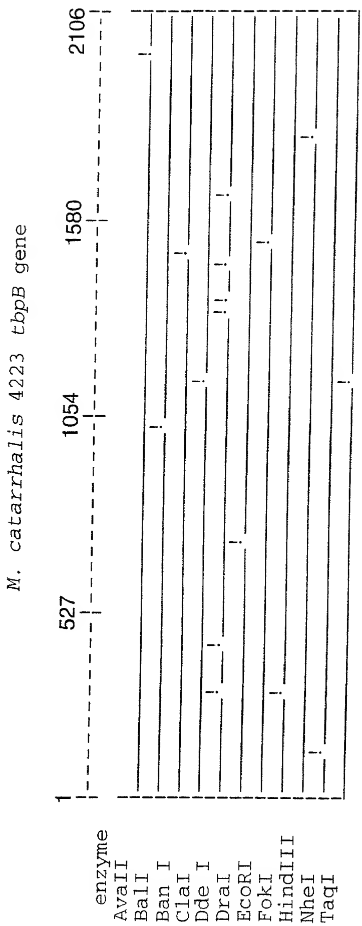

M. catarrhalis 4223 tbpB gene
1
527
1054
1580
2106
enzyme
AvaII
BalI
Ban I
ClaI
Dde I
DraI
EcoRI
FokI
HindIII
NheI
TaqI

FIG.4

FIG.5A

Sequence of *M. catarrhalis* 4223 *tbpA* gene

```
TATTTTGACAAGCTATACACTAAAATCAAAAATTAATCACTTTGGTTGGGTGGTTTTAGCAAGCAAATGGT

TATTTTGGTAAACAATTAAGTTCTTAAAAACGATACACGCTCATAAACAGATGGTTTTTGGCATCTGCAAT

TTGATGCCTGCCTTGTGATTGGTTGGGGTGTATCGGTGTATCAAAGTGCAAAAGCCAACAGGTGGTCATTG

                                27                                  54
ATG AAT CAA TCA AAA CAA AAC AAC AAA TCC AAA AAA TCC AAA CAA GTA TTA AAA
MET Asn Gln Ser Lys Gln Asn Asn Lys Ser Lys Lys Ser Lys Gln Val Leu Lys

                                81                                 108
CTT AGT GCC TTG TCT TTG GGT CTG CTT AAC ATC ACG CAG GTG GCA CTG GCA AAC
Leu Ser Ala Leu Ser Leu Gly Leu Leu Asn Ile Thr Gln Val Ala Leu Ala Asn

                               135                                 162
ACA ACG GCC GAT AAG GCG GAG GCA ACA GAT AAG ACA AAC CTT GTT GTT GTC TTG
Thr Thr Ala Asp Lys Ala Glu Ala Thr Asp Lys Thr Asn Leu Val Val Val Leu

                               189                                 216
GAT GAA ACT GTT GTA ACA GCG AAG AAA AAC GCC CGT AAA GCC AAC GAA GTT ACA
Asp Glu Thr Val Val Thr Ala Lys Lys Asn Ala Arg Lys Ala Asn Glu Val Thr
```

FIG.5B

```
                                                243                                             270
GGG CTT GGT AAG GTG GTC AAA ACT GCC GAG ACC ATC AAT AAA GAA CAA GTG CTA
Gly Leu Gly Lys Val Val Lys Thr Ala Glu Thr Ile Asn Lys Glu Gln Val Leu

                                                297                                             324
AAC ATT CGA GAC TTA ACA CGC TAT GAC CCT GGC ATT GCT GTG GTT GAG CAA GGT
Asn Ile Arg Asp Leu Thr Arg Tyr Asp Pro Gly Ile Ala Val Val Glu Gln Gly

                                                351                                             378
CGT GGG GCA AGC TCA GGC TAT TCT ATT CGT GGT ATG GAT AAA AAT CGT GTG GCG
Arg Gly Ala Ser Ser Gly Tyr Ser Ile Arg Gly MET Asp Lys Asn Arg Val Ala

                                                405                                             432
GTA TTG GTT GAT GGC ATC AAT CAA GCC CAG CAC TAT GCC CTA CAA GGC CCT GTG
Val Leu Val Asp Gly Ile Asn Gln Ala Gln His Tyr Ala Leu Gln Gly Pro Val

                                                459                                             486
GCA GGC AAA AAT TAT GCC GCA GGT GGG GCA ATC AAC GAA ATA GAA TAC GAA AAT
Ala Gly Lys Asn Tyr Ala Ala Gly Gly Ala Ile Asn Glu Ile Glu Tyr Glu Asn

                                                513                                             540
GTC CGC TCC GTT GAG ATT AGT AAA GGT GCA AAT TCA AGT GAA TAC GGC TCT GGG
Val Arg Ser Val Glu Ile Ser Lys Gly Ala Asn Ser Ser Glu Tyr Gly Ser Gly
```

FIG.5C

```
                                        567                                                 594
GCA TTA TCT GGC TCT GTG GCA TTT GTT ACC AAA ACC GCC GAT GAC ATC ATC AAA
Ala Leu Ser Gly Ser Val Ala Phe Val Thr Lys Thr Ala Asp Asp Ile Ile Lys

                                        621                                                 648
GAT GGT AAA GAT TGG GGC GTG CAG ACC AAA ACC GCC TAT GCC AGT AAA AAT AAC
Asp Gly Lys Asp Trp Gly Val Gln Thr Lys Thr Ala Tyr Ala Ser Lys Asn Asn

                                        675                                                 702
GCA TGG GTT AAT TCT GTG GCA GCA GCA GGC AAG GCA GGT TCT TTT AGC GGT CTT
Ala Trp Val Asn Ser Val Ala Ala Ala Gly Lys Ala Gly Ser Phe Ser Gly Leu

                                        729                                                 756
ATC ATC TAC ACC GAC CGC CGT GGT CAA GAA TAC AAG GCA CAT GAT GAT GCC TAT
Ile Ile Tyr Thr Asp Arg Arg Gly Gln Glu Tyr Lys Ala His Asp Asp Ala Tyr

                                        783                                                 810
CAG GGT AGC CAA AGT TTT GAT AGA GCG GTG GCA ACC ACT GAC CCA AAT AAC CGA
Gln Gly Ser Gln Ser Phe Asp Arg Ala Val Ala Thr Thr Asp Pro Asn Asn Arg

                                        837                                                 864
ACA TTT TTA ATA GCA AAT GAA TGT GCC AAT GGT AAT TAT GAG GCG TGT GCT GCT
Thr Phe Leu Ile Ala Asn Glu Cys Ala Asn Gly Asn Tyr Glu Ala Cys Ala Ala

                                        891                                                 918
GGC GGT CAA ACC AAA CTT CAA GCC AAG CCA ACC AAT GTG CGT GAT AAG GTC AAT
Gly Gly Gln Thr Lys Leu Gln Ala Lys Pro Thr Asn Val Arg Asp Lys Val Asn
```

09/142628

FIG.5D

```
                                945                                        972
GTC AAA GAT TAT ACA GGT CCT AAC CGC CTT ATC CCA AAC CCA CTC ACC CAA GAC
Val Lys Asp Tyr Thr Gly Pro Asn Arg Leu Ile Pro Asn Pro Leu Thr Gln Asp

                                999                                       1026
AGC AAA TCC TTA CTG CTT CGC CCA GGT TAT CAG CTA AAC GAT AAG CAC TAT GTC
Ser Lys Ser Leu Leu Leu Arg Pro Gly Tyr Gln Leu Asn Asp Lys His Tyr Val

                               1053                                       1080
GGT GGT GTG TAT GAA ATC ACC AAA CAA AAC TAC GCC ATG CAA GAT AAA ACC GTG
Gly Gly Val Tyr Glu Ile Thr Lys Gln Asn Tyr Ala MET Gln Asp Lys Thr Val

                               1107                                       1134
CCT GCT TAT CTG ACG GTT CAT GAC ATT GAA AAA TCA AGG CTC AGC AAC CAT GCC
Pro Ala Tyr Leu Thr Val His Asp Ile Glu Lys Ser Arg Leu Ser Asn His Ala

                               1161                                       1188
CAA GCC AAT GGC TAT TAT CAA GGC AAT AAT CTT GGT GAA CGC ATT CGT GAT ACC
Gln Ala Asn Gly Tyr Tyr Gln Gly Asn Asn Leu Gly Glu Arg Ile Arg Asp Thr

                               1215                                       1242
ATT GGG CCA GAT TCA GGT TAT GGC ATC AAC TAT GCT CAT GGC GTA TTT TAT GAT
Ile Gly Pro Asp Ser Gly Tyr Gly Ile Asn Tyr Ala His Gly Val Phe Tyr Asp
```

09/142628

FIG.5E

```
                                                    1269                                              1296
GAA AAA CAC CAA AAA GAC CGC CTA GGG CTT GAA TAT GTT TAT GAC AGC AAA GGT
Glu Lys His Gln Lys Asp Arg Leu Gly Leu Glu Tyr Val Tyr Asp Ser Lys Gly

                                                    1323                                              1350
GAA AAT AAA TGG TTT GAT GAT GTG CGT GTG TCT TAT GAT AAG CAA GAC ATT ACG
Glu Asn Lys Trp Phe Asp Asp Val Arg Val Ser Tyr Asp Lys Gln Asp Ile Thr

                                                    1377                                              1404
CTA CGC AGC CAG CTG ACC AAC ACG CAC TGT TCA ACC TAT CCG CAC ATT GAC AAA
Leu Arg Ser Gln Leu Thr Asn Thr His Cys Ser Thr Tyr Pro His Ile Asp Lys

                                                    1431                                              1458
AAT TGT ACG CCT GAT GTC AAT AAA CCT TTT TCG GTA AAA GAG GTG GAT AAC AAT
Asn Cys Thr Pro Asp Val Asn Lys Pro Phe Ser Val Lys Glu Val Asp Asn Asn

                                                    1485                                              1512
GCC TAC AAA GAA CAG CAC AAT TTA ATC AAA GCC GTC TTT AAC AAA AAA ATG GCG
Ala Tyr Lys Glu Gln His Asn Leu Ile Lys Ala Val Phe Asn Lys Lys MET Ala

                                                    1539                                              1566
TTG GGC AGT ACG CAT CAT CAC ATC AAC CTG CAA GTT GGC TAT GAT AAA TTC AAT
Leu Gly Ser Thr His His His Ile Asn Leu Gln Val Gly Tyr Asp Lys Phe Asn

                                                    1593                                              1620
TCA AGC CTG AGC CGT GAA GAT TAT CGT TTG GCA ACC CAT CAG TCT TAT CAA AAA
Ser Ser Leu Ser Arg Glu Asp Tyr Arg Leu Ala Thr His Gln Ser Tyr Gln Lys
```

FIG.5F

```
                                             1647                                            1674
CTT GAT TAC ACC CCA CCA AGT AAC CCT TTG CCA GAT AAG TTT AAG CCC ATT TTA
Leu Asp Tyr Thr Pro Pro Ser Asn Pro Leu Pro Asp Lys Phe Lys Pro Ile Leu

                                             1701                                            1728
GGT TCA AAC AAC AAA CCC ATT TGC CTT GAT GCT TAT GGT TAT GGT CAT GAC CAT
Gly Ser Asn Asn Lys Pro Ile Cys Leu Asp Ala Tyr Gly Tyr Gly His Asp His

                                             1755                                            1782
CCA CAG GCT TGT AAC GCC AAA AAC AGC ACT TAT CAA AAT TTT GCC ATC AAA AAA
Pro Gln Ala Cys Asn Ala Lys Asn Ser Thr Tyr Gln Asn Phe Ala Ile Lys Lys

                                             1809                                            1836
GGC ATA GAG CAA TAC AAC CAA AAA ACC AAT ACC GAT AAG ATT GAT TAT CAA GCC
Gly Ile Glu Gln Tyr Asn Gln Lys Thr Asn Thr Asp Lys Ile Asp Tyr Gln Ala

                                             1863                                            1890
ATC ATT GAC CAA TAT GAT AAA CAA AAC CCC AAC AGC ACC CTA AAA CCC TTT GAG
Ile Ile Asp Gln Tyr Asp Lys Gln Asn Pro Asn Ser Thr Leu Lys Pro Phe Glu

                                             1917                                            1944
AAA ATC AAA CAA AGT TTG GGG CAA GAA AAA TAC AAC AAG ATA GAC GAA CTT GGC
Lys Ile Lys Gln Ser Leu Gly Gln Glu Lys Tyr Asn Lys Ile Asp Glu Leu Gly
```

09/142628

FIG.5G

```
                                    1971                                        1998
TTT AAA GCT TAT AAA GAT TTA CGC AAC GAA TGG GCG GGT TGG ACT AAT GAC AAC
Phe Lys Ala Tyr Lys Asp Leu Arg Asn Glu Trp Ala Gly Trp Thr Asn Asp Asn

                                    2025                                        2052
AGC CAA CAA AAT GCC AAT AAA GGC ACG GAT AAT ATC TAT CAG CCA AAT CAA GCA
Ser Gln Gln Asn Ala Asn Lys Gly Thr Asp Asn Ile Tyr Gln Pro Asn Gln Ala

                                    2079                                        2106
ACT GTG GTC AAA GAT GAC AAA TGT AAA TAT AGC GAG ACC AAC AGC TAT GCT GAT
Thr Val Val Lys Asp Asp Lys Cys Lys Tyr Ser Glu Thr Asn Ser Tyr Ala Asp

                                    2133                                        2160
TGC TCA ACC ACT CGC CAC ATC AGT GGT GAT AAT TAT TTC ATC GCT TTA AAA GAC
Cys Ser Thr Thr Arg His Ile Ser Gly Asp Asn Tyr Phe Ile Ala Leu Lys Asp

                                    2187                                        2214
AAC ATG ACC ATC AAT AAA TAT GTT GAT TTG GGG CTG GGT GCT CGC TAT GAC AGA
Asn MET Thr Ile Asn Lys Tyr Val Asp Leu Gly Leu Gly Ala Arg Tyr Asp Arg

                                    2241                                        2268
ATC AAA CAC AAA TCT GAT GTG CCT TTG GTA GAC AAC AGT GCC AGC AAC CAG CTG
Ile Lys His Lys Ser Asp Val Pro Leu Val Asp Asn Ser Ala Ser Asn Gln Leu
```

09/142628

FIG.5H

```
                                        2295                                        2322
TCT TGG AAT TTT GGC GTG GTC GTC AAG CCC ACC AAT TGG CTG GAC ATC GCT TAT
Ser Trp Asn Phe Gly Val Val Val Lys Pro Thr Asn Trp Leu Asp Ile Ala Tyr

                                        2349                                        2376
AGA AGC TCG CAA GGC TTT CGC ATG CCA AGT TTT TCT GAA ATG TAT GGC GAA CGC
Arg Ser Ser Gln Gly Phe Arg MET Pro Ser Phe Ser Glu MET Tyr Gly Glu Arg

                                        2403                                        2430
TTT GGC GTA ACC ATC GGT AAA GGC ACG CAA CAT GGC TGT AAG GGT CTT TAT TAC
Phe Gly Val Thr Ile Gly Lys Gly Thr Gln His Gly Cys Lys Gly Leu Tyr Tyr

                                        2457                                        2484
ATT TGT CAG CAG ACT GTC CAT CAA ACC AAG CTA AAA CCT GAA AAA TCC TTT AAC
Ile Cys Gln Gln Thr Val His Gln Thr Lys Leu Lys Pro Glu Lys Ser Phe Asn

                                        2511                                        2538
CAA GAA ATC GGA GCG ACT TTA CAT AAC CAC TTA GGC AGT CTT GAG GTT AGT TAT
Gln Glu Ile Gly Ala Thr Leu His Asn His Leu Gly Ser Leu Glu Val Ser Tyr

                                        2565                                        2592
TTT AAA AAT CGC TAT ACC GAT TTG ATT GTT GGT AAA AGT GAA GAG ATT AGA ACC
Phe Lys Asn Arg Tyr Thr Asp Leu Ile Val Gly Lys Ser Glu Glu Ile Arg Thr

                                        2619                                        2646
CTA ACC CAA GGT GAT AAT GCA GGC AAA CAG CGT GGT AAA GGT GAT TTG GGC TTT
Leu Thr Gln Gly Asp Asn Ala Gly Lys Gln Arg Gly Lys Gly Asp Leu Gly Phe
```

FIG.5I

```
                                    2673                                 2700
CAT AAT GGA CAA GAT GCT GAT TTG ACA GGC ATT AAC ATT CTT GGC AGA CTT GAC
His Asn Gly Gln Asp Ala Asp Leu Thr Gly Ile Asn Ile Leu Gly Arg Leu Asp

                                    2727                                 2754
CTA AAC GCT GTC AAT AGT CGC CTT CCC TAT GGA TTA TAC TCA ACA CTG GCT TAT
Leu Asn Ala Val Asn Ser Arg Leu Pro Tyr Gly Leu Tyr Ser Thr Leu Ala Tyr

                                    2781                                 2808
AAC AAA GTT GAT GTT AAA GGA AAA ACC TTA AAC CCA ACT TTG GCA GGA ACA AAC
Asn Lys Val Asp Val Lys Gly Lys Thr Leu Asn Pro Thr Leu Ala Gly Thr Asn

                                    2835                                 2862
ATA CTG TTT GAT GCC ATC CAG CCA TCT CGT TAT GTG GTG GGG CTT GGC TAT GAT
Ile Leu Phe Asp Ala Ile Gln Pro Ser Arg Tyr Val Val Gly Leu Gly Tyr Asp

                                    2889                                 2916
GCC CCA AGC CAA AAA TGG GGA GCA AAC GCC ATA TTT ACC CAT TCT GAT GCC AAA
Ala Pro Ser Gln Lys Trp Gly Ala Asn Ala Ile Phe Thr His Ser Asp Ala Lys

                                    2943                                 2970
AAT CCA AGC GAG CTT TTG GCA GAT AAG AAC TTA GGT AAT GGC AAC ATT CAA ACA
Asn Pro Ser Glu Leu Leu Ala Asp Lys Asn Leu Gly Asn Gly Asn Ile Gln Thr
```

FIG.5J

```
                                            2997                                              3024
AAA CAA GCC ACC AAA GCA AAA TCC ACG CCG TGG CAA ACA CTT GAT TTG TCA GGT
Lys Gln Ala Thr Lys Ala Lys Ser Thr Pro Trp Gln Thr Leu Asp Leu Ser Gly

                                            3051                                              3078
TAT GTA AAC ATA AAA GAT AAT TTT ACC TTG CGT GCT GGC GTG TAC AAT GTA TTT
Tyr Val Asn Ile Lys Asp Asn Phe Thr Leu Arg Ala Gly Val Tyr Asn Val Phe

                                            3105                                              3132
AAT ACC TAT TAC ACC ACT TGG GAG GCT TTA CGC CAA ACA GCA GAA GGG GCG GTC
Asn Thr Tyr Tyr Thr Thr Trp Glu Ala Leu Arg Gln Thr Ala Glu Gly Ala Val

                                            3159                                              3186
AAT CAG CAT ACA GGA CTG AGC CAA GAT AAG CAT TAT GGT CGC TAT GCC GCT CCT
Asn Gln His Thr Gly Leu Ser Gln Asp Lys His Tyr Gly Arg Tyr Ala Ala Pro

                                            3213
GGA CGC AAT TAC CAA TTG GCA CTT GAA ATG AAG TTT TAA
Gly Arg Asn Tyr Gln Leu Ala Leu Glu MET Lys Phe  .
```

FIG.6A

Sequence of *M. catarrhalis* 4223 *tbpB* gene

```
GTAAATTTGCCGTATTTTGTCTATCATAAATGCATTTATCAAATGCTCAAATAAATACGCCAAATGCACAT

TGTCAGCATGCCAAAATAGGCATCAACAGACTTTTTAGATAATACCATCAACCCATCAGAGGATTATTTT

                                27                                  54
ATG AAA CAC ATT CCT TTA ACC ACA CTG TGT GTG GCA ATC TCT GCC GTC TTA TTA
MET Lys His Ile Pro Leu Thr Thr Leu Cys Val Ala Ile Ser Ala Val Leu Leu

                                81                                 108
ACC GCT TGT GGT GGC AGT GGT GGT TCA AAT CCA CCT GCT CCT ACG CCC ATT CCA
Thr Ala Cys Gly Gly Ser Gly Gly Ser Asn Pro Pro Ala Pro Thr Pro Ile Pro

                               135                                 162
AAT GCT AGC GGT TCA GGT AAT ACT GGC AAC ACT GGT AAT GCT GGC GGT ACT GAT
Asn Ala Ser Gly Ser Gly Asn Thr Gly Asn Thr Gly Asn Ala Gly Gly Thr Asp

                               189                                 216
AAT ACA GCC AAT GCA GGT AAT ACA GGC GGT ACA AAC TCT GGT ACA GGC AGT GCC
Asn Thr Ala Asn Ala Gly Asn Thr Gly Gly Thr Asn Ser Gly Thr Gly Ser Ala

                               243                                 270
AAC ACA CCA GAG CCA AAA TAT CAA GAT GTA CCA ACT GAG AAA AAT GAA AAA GAT
Asn Thr Pro Glu Pro Lys Tyr Gln Asp Val Pro Thr Glu Lys Asn Glu Lys Asp
```

FIG.6B

```
                                                297                                        324
AAA GTT TCA TCC ATT CAA GAA CCT GCC ATG GGT TAT GGC ATG GCT TTG AGT AAA
Lys Val Ser Ser Ile Gln Glu Pro Ala MET Gly Tyr Gly MET Ala Leu Ser Lys

                                                351                                        378
ATT AAT CTA CAC AAC CGA CAA GAC ACG CCA TTA GAT GAA AAA AAT ATC ATT ACC
Ile Asn Leu His Asn Arg Gln Asp Thr Pro Leu Asp Glu Lys Asn Ile Ile Thr

                                                405                                        432
TTA GAC GGT AAA AAA CAA GTT GCA GAA GGT AAA AAA TCG CCA TTG CCA TTT TCG
Leu Asp Gly Lys Lys Gln Val Ala Glu Gly Lys Lys Ser Pro Leu Pro Phe Ser

                                                459                                        486
TTA GAT GTA GAA AAT AAA TTG CTT GAT GGC TAT ATA GCA AAA ATG AAT GTA GCG
Leu Asp Val Glu Asn Lys Leu Leu Asp Gly Tyr Ile Ala Lys MET Asn Val Ala

                                                513                                        540
GAT AAA AAT GCC ATT GGT GAC AGA ATT AAG AAA GGT AAT AAA GAA ATC TCC GAT
Asp Lys Asn Ala Ile Gly Asp Arg Ile Lys Lys Gly Asn Lys Glu Ile Ser Asp

                                                567                                        594
GAA GAA CTT GCC AAA CAA ATC AAA GAA GCT GTG CGT AAA AGC CAT GAG TTT CAG
Glu Glu Leu Ala Lys Gln Ile Lys Glu Ala Val Arg Lys Ser His Glu Phe Gln
```

FIG.6C

```
                                                621                                                  648
CAA GTA TTA TCA TCA CTG GAA AAC AAA ATT TTT CAT TCA AAT GAC GGA ACA ACC
Gln Val Leu Ser Ser Leu Glu Asn Lys Ile Phe His Ser Asn Asp Gly Thr Thr

                                                675                                                  702
AAA GCA ACC ACA CGA GAT TTA AAA TAT GTT GAT TAT GGT TAC TAC TTG GCG AAT
Lys Ala Thr Thr Arg Asp Leu Lys Tyr Val Asp Tyr Gly Tyr Tyr Leu Ala Asn

                                                729                                                  756
GAT GGC AAT TAT CTA ACC GTC AAA ACA GAC AAA CTT TGG AAT TTA GGC CCT GTG
Asp Gly Asn Tyr Leu Thr Val Lys Thr Asp Lys Leu Trp Asn Leu Gly Pro Val

                                                783                                                  810
GGT GGT GTG TTT TAT AAT GGC ACA ACG ACC GCC AAA GAG TTG CCC ACA CAA GAT
Gly Gly Val Phe Tyr Asn Gly Thr Thr Thr Ala Lys Glu Leu Pro Thr Gln Asp

                                                837                                                  864
GCG GTC AAA TAT AAA GGA CAT TGG GAC TTT ATG ACC GAT GTT GCC AAC AGA AGA
Ala Val Lys Tyr Lys Gly His Trp Asp Phe MET Thr Asp Val Ala Asn Arg Arg

                                                891                                                  918
AAC CGA TTT AGC GAA GTG AAA GAA AAC TCT CAA GCA GGC TGG TAT TAT GGA GCA
Asn Arg Phe Ser Glu Val Lys Glu Asn Ser Gln Ala Gly Trp Tyr Tyr Gly Ala
```

FIG.6D

```
                                945                                      972
TCT TCA AAA GAT GAA TAC AAC CGC TTA TTA ACT AAA GAA GAC TCT GCC CCT GAT
Ser Ser Lys Asp Glu Tyr Asn Arg Leu Leu Thr Lys Glu Asp Ser Ala Pro Asp

                                999                                     1026
GGT CAT AGC GGT GAA TAT GGC CAT AGC AGT GAG TTT ACT GTT AAT TTT AAG GAA
Gly His Ser Gly Glu Tyr Gly His Ser Ser Glu Phe Thr Val Asn Phe Lys Glu

                               1053                                     1080
AAA AAA TTA ACA GGT AAG CTG TTT AGT AAC CTA CAA GAC CGC CAT AAG GGC AAT
Lys Lys Leu Thr Gly Lys Leu Phe Ser Asn Leu Gln Asp Arg His Lys Gly Asn

                               1107                                     1134
GTT ACA AAA ACC GAA CGC TAT GAC ATC GAT GCC AAT ATC CAC GGC AAC CGC TTC
Val Thr Lys Thr Glu Arg Tyr Asp Ile Asp Ala Asn Ile His Gly Asn Arg Phe

                               1161                                     1188
CGT GGC AGT GCC ACC GCA AGC AAT AAA AAT GAC ACA AGC AAA CAC CCC TTT ACC
Arg Gly Ser Ala Thr Ala Ser Asn Lys Asn Asp Thr Ser Lys His Pro Phe Thr

                               1215                                     1242
AGT GAT GCC AAC AAT AGG CTA GAA GGT GGT TTT TAT GGG CCA AAA GGC GAG GAG
Ser Asp Ala Asn Asn Arg Leu Glu Gly Gly Phe Tyr Gly Pro Lys Gly Glu Glu
```

09/142628

FIG.6E

```
                                                  1269                                              1296
CTG GCA GGT AAA TTC TTA ACC AAT GAC AAC AAA CTC TTT GGC GTC TTT GGT GCT
Leu Ala Gly Lys Phe Leu Thr Asn Asp Asn Lys Leu Phe Gly Val Phe Gly Ala

                                                  1323                                              1350
AAA CGA GAG AGT AAA GCT GAG GAA AAA ACC GAA GCC ATC TTA GAT GCC TAT GCA
Lys Arg Glu Ser Lys Ala Glu Glu Lys Thr Glu Ala Ile Leu Asp Ala Tyr Ala

                                                  1377                                              1404
CTT GGG ACA TTT AAT ACA AGT AAC GCA ACC ACA TTC ACC CCA TTT ACC GAA AAA
Leu Gly Thr Phe Asn Thr Ser Asn Ala Thr Thr Phe Thr Pro Phe Thr Glu Lys

                                                  1431                                              1458
CAA CTG GAT AAC TTT GGC AAT GCC AAA AAA TTG GTC TTA GGT TCT ACC GTC ATT
Gln Leu Asp Asn Phe Gly Asn Ala Lys Lys Leu Val Leu Gly Ser Thr Val Ile

                                                  1485                                              1512
GAT TTG GTG CCT ACT GAT GCC ACC AAA AAT GAA TTC ACC AAA GAC AAG CCA GAG
Asp Leu Val Pro Thr Asp Ala Thr Lys Asn Glu Phe Thr Lys Asp Lys Pro Glu

                                                  1539                                              1566
TCT GCC ACA AAC GAA GCG GGC GAG ACT TTG ATG GTG AAT GAT GAA GTT AGC GTC
Ser Ala Thr Asn Glu Ala Gly Glu Thr Leu MET Val Asn Asp Glu Val Ser Val
```

FIG.6F

```
                                    1593                                       1620
AAA ACC TAT GGC AAA AAC TTT GAA TAC CTA AAA TTT GGT GAG CTT AGT ATC GGT
Lys Thr Tyr Gly Lys Asn Phe Glu Tyr Leu Lys Phe Gly Glu Leu Ser Ile Gly

                                    1647                                       1674
GGT AGC CAT AGC GTC TTT TTA CAA GGC GAA CGC ACC GCT ACC ACA GGC GAG AAA
Gly Ser His Ser Val Phe Leu Gln Gly Glu Arg Thr Ala Thr Thr Gly Glu Lys

                                    1701                                       1728
GCC GTA CCA ACC ACA GGC ACA GCC AAA TAT TTG GGG AAC TGG GTA GGA TAC ATC
Ala Val Pro Thr Thr Gly Thr Ala Lys Tyr Leu Gly Asn Trp Val Gly Tyr Ile

                                    1755                                       1782
ACA GGA AAG GAC ACA GGA ACG GGC ACA GGA AAA AGC TTT ACC GAT GCC CAA GAT
Thr Gly Lys Asp Thr Gly Thr Gly Thr Gly Lys Ser Phe Thr Asp Ala Gln Asp

                                    1809                                       1836
GTT GCT GAT TTT GAC ATT GAT TTT GGA AAT AAA TCA GTC AGC GGT AAA CTT ATC
Val Ala Asp Phe Asp Ile Asp Phe Gly Asn Lys Ser Val Ser Gly Lys Leu Ile

                                    1863                                       1890
ACC AAA GGC CGC CAA GAC CCT GTA TTT AGC ATC ACA GGT CAA ATC GCA GGC AAT
Thr Lys Gly Arg Gln Asp Pro Val Phe Ser Ile Thr Gly Gln Ile Ala Gly Asn
```

FIG.6G

```
                                1917                                          1944
GGC TGG ACA GGG ACA GCC AGC ACC ACC AAA GCG GAC GCA GGA GGC TAC AAG ATA
Gly Trp Thr Gly Thr Ala Ser Thr Thr Lys Ala Asp Ala Gly Gly Tyr Lys Ile

                                1971                                          1998
GAT TCT AGC AGT ACA GGC AAA TCC ATC GCC ATC AAA GAT GCC AAT GTT ACA GGG
Asp Ser Ser Ser Thr Gly Lys Ser Ile Ala Ile Lys Asp Ala Asn Val Thr Gly

                                2025                                          2052
GGC TTT TAT GGT CCA AAT GCA AAC GAG ATG GGC GGG TCA TTT ACA CAC AAC GCC
Gly Phe Tyr Gly Pro Asn Ala Asn Glu MET Gly Gly Ser Phe Thr His Asn Ala

                                2079                                          2106
GAT GAC AGC AAA GCC TCT GTG GTC TTT GGC ACA AAA AGA CAA CAA GAA GTT AAG
Asp Asp Ser Lys Ala Ser Val Val Phe Gly Thr Lys Arg Gln Gln Glu Val Lys
```

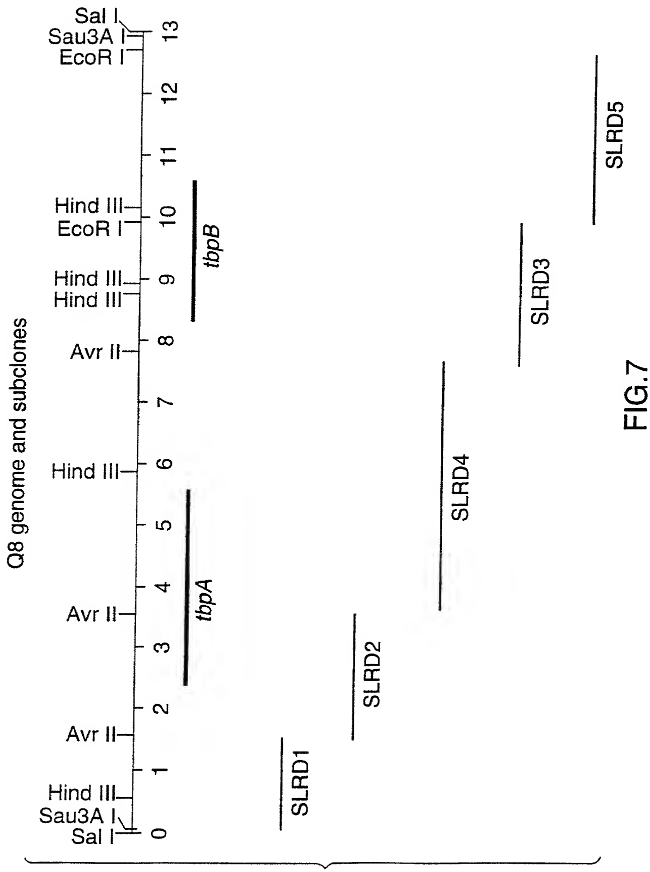

Q8 genome and subclones
Sal I
Sau3A I
EcoR I
Hind III
EcoR I
Hind III
Hind III
Avr II
Hind III
Avr II
Avr II
Hind III
Sau3A I
Sal I
0
1
2
3
4
5
6
7
8
9
10
11
12
13
tbpB
tbpA
SLRD1
SLRD2
SLRD4
SLRD3
SLRD5

FIG.7

09/142628

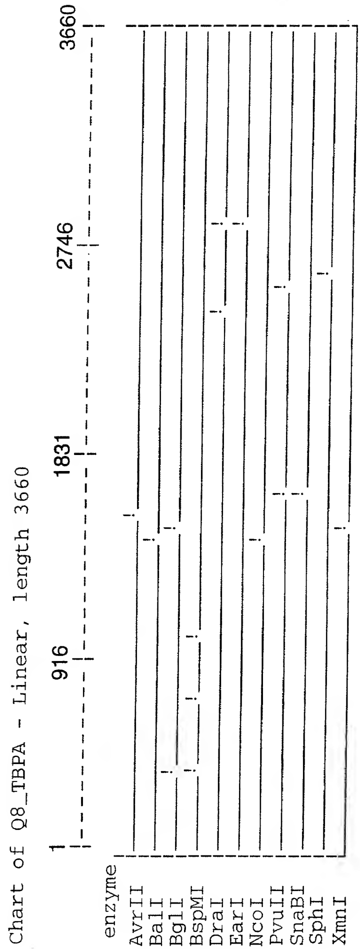


FIG.8

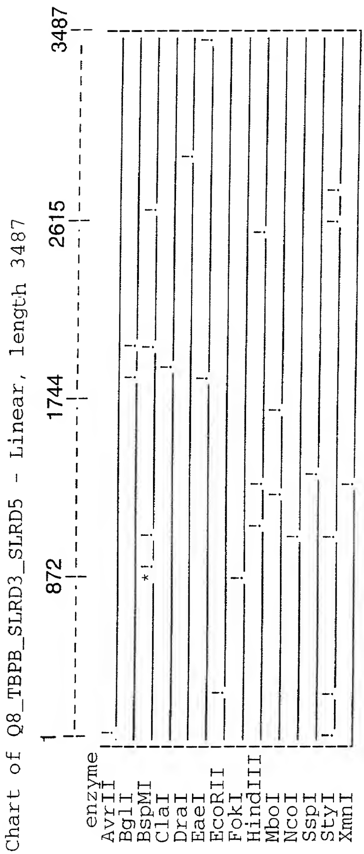


Chart of Q8_TBPB_SLRD3_SLRD5 - Linear, length 3487
1
872
1744
2615
3487
enzyme
AvrII
BglI
BspMI
ClaI
DraI
EaeI
EcoRII
FokI
HindIII
MboI
NcoI
SspI
StyI
XmnI

FIG.9

FIG.10A

Q8 tbpA gene sequence

```
A A T T G A T A C A A A A T G G T T T G T A T T A T C A C T
                  10                  20                  30
                                          T G T A T T T G T A T T A T A A T T T T A C T T A T T T T T
                                                            40                  50                  60

A C A A A C T A T A C A C T A A A A T C A A A A A T T A A T
                  70                  80                  90
                                          C A C T T T G G T T G G G T G G T T T T A G C A A G C A A A
                                                           100                 110                 120

T G G T T A T T T T G G T A A A C A A T T A A G T T C T T A
                 130                 140                 150
                                          A A A A C G A T A C A C G C T C A T A A A C A G A T G G T T
                                                           160                 170                 180

T T T G G C A T C T T C A A T T T G A T G C C T G C C T T G
                 190                 200                 210
                                          T G A T T G G T T G G G G G T G T A T T G A T G T A T C C A
                                                           220                 230                 240

                                                        MET
A G T A C A A A A G C C A A C A G G T G G T C A T T G A T G
                 250                 260                 270
```

09/142628

FIG.10B

```
                              ASN GLN SER LYS LYS SER LYS LYS SER LYS
                              A A T C A A T C C A A A A A A T C C A A A A A A T C C A A A
                                                  280                 290                 300

GLN VAL LEU LYS LEU SER ALA LEU SER LEU
C A A G T A T T A A A A C T T A G T G C C T T G T C T T T G
                310                 320                 330
                              GLY LEU LEU ASN ILE THR GLN VAL ALA LEU
                              G G T C T G C T T A A C A T C A C G C A G G T G G C A C T G
                                                  340                 350                 360

ALA ASN THR THR ALA ASP LYS ALA GLU ALA
G C A A A C A C A A C G G C C G A T A A G G C G G A G G C A
                370                 380                 390
                              THR ASP LYS THR ASN LEU VAL VAL VAL LEU
                              A C A G A T A A G A C A A A C C T T G T T G T T G T C T T G
                                                  400                 410                 420

ASP GLU THR VAL VAL THR ALA LYS LYS ASN
G A T G A A A C T G T T G T A A C A G C G A A G A A A A A C
                430                 440                 450
                              ALA ARG LYS ALA ASN GLU VAL THR GLY LEU
                              G C C C G T A A A G C C A A C G A A G T T A C A G G G C T T
                                                  460                 470                 480
```

FIG.10C

```
GLY LYS VAL VAL LYS THR ALA GLU THR ILE
GGTAAGGTGGTCAAAACTGCCGAGACCATC
        490       500       510
                              ASN LYS GLU GLN VAL LEU ASN ILE ARG ASP
                              AATAAAGAACAAGTGCTAAACATTCGAGAC
                                      520       530       540

LEU THR ARG TYR ASP PRO GLY ILE ALA VAL
TTAACACGCTATGACCCTGGCATTGCTGTG
        550       560       570
                              VAL GLU GLN GLY ARG GLY ALA SER SER GLY
                              GTTGAGCAAGGTCGTGGGGCAAGCTCAGGC
                                      580       590       600

TYR SER ILE ARG GLY MET ASP LYS ASN ARG
TATTCTATTCGTGGTATGGATAAAAATCGT
        610       620       630
                              VAL ALA VAL LEU VAL ASP GLY ILE ASN GLN
                              GTGGCGGTATTGGTTGATGGCATCAATCAA
                                      640       650       660

ALA GLN HIS TYR ALA LEU GLN GLY PRO VAL
GCCCAGCACTATGCCCTACAAGGCCCTGTG
        670       680       690
```

FIG.10D

```
ALA GLY LYS ASN TYR ALA ALA GLY GLY ALA
G C A G G C A A A A A T T A T G C C G C A G G T G G G G C A
                    700                 710                 720

ILE ASN GLU ILE GLU TYR GLU ASN VAL ARG
A T C A A C G A A A T A G A A T A C G A A A A T G T C C G C
                    730                 740                 750

SER VAL GLU ILE SER LYS GLY ALA ASN SER
T C C G T T G A G A T T A G T A A A G G T G C A A A T T C A
                    760                 770                 780

SER GLU TYR GLY SER GLY ALA LEU SER GLY
A G T G A A T A C G G C T C T G G G G C A T T A T C T G G C
                    790                 800                 810

SER VAL ALA PHE VAL THR LYS THR ALA ASP
T C T G T G G C A T T T G T T A C C A A A A C C G C C G A T
                    820                 830                 840

ASP ILE ILE LYS ASP GLY LYS ASP TRP GLY
G A C A T C A T C A A A G A T G G T A A A G A T T G G G G C
                    850                 860                 870

VAL GLN THR LYS THR ALA TYR ALA SER LYS
G T G C A G A C C A A A A C C G C C T A T G C C A G T A A A
                    880                 890                 900
```

FIG.10E

```
ASN   ASN   ALA   TRP   VAL   ASN   SER   VAL   ALA   ALA
A A T A A C G C A T G G G T T A A T T C T G T G G C A G C A
              910                   920                   930
                                          ALA   GLY   LYS   ALA   GLY   SER   PHE   SER   GLY   LEU
                                          G C A G G C A A G G C A G G T T C T T T T A G C G G T C T T
                                                        940                   950                   960

ILE   ILE   TYR   THR   ASP   ARG   ARG   GLY   GLN   GLU
A T C A T C T A C A C C G A C C G C C G T G G T C A A G A A
              970                   980                   990
                                          TYR   LYS   ALA   HIS   ASP   ASP   ALA   TYR   GLN   GLY
                                          T A C A A G G C A C A T G A T G A T G C C T A T C A G G G T
                                                       1000                  1010                  1020

SER   GLN   SER   PHE   ASP   ARG   ALA   VAL   ALA   THR
A G C C A A A G T T T T G A T A G A G C G G T G G C A A C C
             1030                  1040                  1050
                                          THR   ASP   PRO   ASN   ASN   PRO   LYS   PHE   LEU   ILE
                                          A C T G A C C C A A A T A A C C C A A A A T T T T T A A T A
                                                       1060                  1070                  1080

ALA   ASN   GLU   CYS   ALA   ASN   GLY   ASN   TYR   GLU
G C A A A T G A A T G T G C C A A T G G T A A T T A T G A G
             1090                  1100                  1110
```

FIG.10F

```
                                   ALA  CYS  ALA  ALA  GLY  GLY  GLN  THR  LYS  LEU
                                   G C G T G T G C T G C T G G C G G T C A A A C C A A A C T C
                                             1120                1130                1140
GLN  ALA  LYS  PRO  THR  ASN  VAL  ARG  ASP  LYS
C A A G C T A A G C C A A C C A A T G T G C G T G A T A A G
          1150                1160                1170
                                   VAL  ASN  VAL  LYS  ASP  TYR  THR  GLY  PRO  ASN
                                   G T C A A T G T C A A A G A T T A T A C A G G T C C T A A C
                                             1180                1190                1200
ARG  LEU  ILE  PRO  ASN  PRO  LEU  THR  GLN  ASP
C G C C T T A T C C C A A A C C C A C T C A C C C A A G A C
          1210                1220                1230
                                   SER  LYS  SER  LEU  LEU  LEU  ARG  PRO  GLY  TYR
                                   A G C A A A T C C T T A C T G C T T C G C C C A G G T T A T
                                             1240                1250                1260
GLN  LEU  ASN  ASP  LYS  HIS  TYR  VAL  GLY  GLY
C A G C T A A A C G A T A A G C A C T A T G T C G G T G G T
          1270                1280                1290
                                   VAL  TYR  GLU  ILE  THR  LYS  GLN  ASN  TYR  ALA
                                   G T G T A T G A A A T C A C C A A A C A A A A C T A C G C C
                                             1300                1310                1320
```

09/142628

FIG.10G

MET GLN ASP LYS THR VAL PRO ALA TYR LEU
A T G C A A G A T A A A A C C G T G C C T G C T T A T C T G
1330 1340 1350

THR VAL HIS ASP ILE GLU LYS SER ARG LEU
A C G G T T C A T G A C A T T G A A A A A T C A A G G C T C
1360 1370 1380

SER ASN HIS GLY GLN ALA ASN GLY TYR TYR
A G C A A C C A T G G C C A A G C C A A T G G C T A T T A T
1390 1400 1410

GLN GLY ASN ASN LEU GLY GLU ARG ILE ARG
C A A G G C A A T A A C C T T G G T G A A C G C A T T C G T
1420 1430 1440

ASP ALA ILE GLY ALA ASN SER GLY TYR GLY
G A T G C C A T T G G G G C A A A T T C A G G T T A T G G C
1450 1460 1470

ILE ASN TYR ALA HIS GLY VAL PHE TYR ASP
A T C A A C T A T G C T C A T G G C G T A T T T T A T G A C
1480 1490 1500

GLU LYS HIS GLN LYS ASP ARG LEU GLY LEU
G A A A A A C A C C A A A A A G A C C G C C T A G G G C T T
1510 1520 1530

09/142628

FIG.10H

```
GLU TYR VAL TYR ASP SER LYS GLY GLU ASN
GAATATGTTTATGACAGCAAAGGTGAAAAT
1540 1550 1560

LYS TRP PHE ASP ASP VAL ARG VAL SER TYR
AAATGGTTTGATGATGTGCGTGTGTCTTAT
1570 1580 1590

ASP LYS GLN ASP ILE THR LEU ARG SER GLN
GACAAGCAAGACATTACGCTACGTAGCCAG
1600 1610 1620

LEU THR ASN THR HIS CYS SER THR TYR PRO
CTGACCAACACGCACTGTTCAACCTATCCG
1630 1640 1650

HIS ILE ASP LYS ASN CYS THR PRO ASP VAL
CACATTGACAAAAATTGTACGCCTGATGTC
1660 1670 1680

ASN LYS PRO PHE SER VAL LYS GLU VAL ASP
AATAAACCTTTTTCGGTAAAAGAGGTGGAT
1690 1700 1710

ASN ASN ALA TYR LYS GLU GLN HIS ASN LEU
AACAATGCCTACAAAGAACAGCACAATTTA
1720 1730 1740
```

09/142628

FIG.10I

```
ILE LYS ALA VAL PHE ASN LYS LYS MET ALA
A T C A A A G C C G T C T T T A A C A A A A A A A T G G C A
              1750                1760              1770
LEU GLY ASN THR HIS HIS HIS ILE ASN LEU
T T G G G C A A T A C G C A T C A T C A C A T C A A T C T G
              1780                1790              1800

GLN VAL GLY TYR ASP LYS PHE ASN SER SER
C A A G T T G G C T A T G A T A A A T T C A A T T C A A G C
              1810                1820              1830
LEU SER ARG GLU ASP TYR ARG LEU ALA THR
C T T A G C C G T G A A G A T T A T C G T T T G G C A A C C
              1840                1850              1860

HIS GLN SER TYR GLN LYS LEU ASP TYR THR
C A T C A A T C T T A T C A A A A A C T T G A T T A C A C C
              1870                1880              1890
PRO PRO SER ASN PRO LEU PRO ASP LYS PHE
C C A C C A A G T A A C C C T T T G C C A G A T A A G T T T
              1900                1910              1920

LYS PRO ILE LEU GLY SER ASN ASN ARG PRO
A A G C C C A T T T T A G G T T C A A A C A A C A G A C C C
              1930                1940              1950
```

FIG.10J

```
                              ILE CYS LEU ASP ALA TYR GLY TYR GLY HIS
                              A T T T G C C T T G A T G C T T A T G G T T A T G G T C A T
                                                1960                1970                1980
ASP HIS PRO GLN ALA CYS ASN ALA LYS ASN
G A C C A T C C A C A G G C T T G T A A C G C C A A A A A C
              1990                2000                2010
                              SER THR TYR GLN ASN PHE ALA ILE LYS LYS
                              A G C A C T T A T C A A A A C T T T G C C A T C A A A A A A
                                                2020                2030                2040
GLY ILE GLU GLN TYR ASN GLN THR ASN THR
G G C A T A G A G C A A T A C A A C C A A A C C A A T A C C
              2050                2060                2070
                              ASP LYS ILE ASP TYR GLN ALA VAL ILE ASP
                              G A T A A G A T T G A T T A T C A A G C C G T C A T T G A C
                                                2080                2090                2100
GLN TYR ASP LYS GLN ASN PRO ASN SER THR
C A A T A T G A T A A A C A A A A C C C C A A C A G C A C C
              2110                2120                2130
                              LEU LYS PRO PHE GLU LYS ILE LYS GLN SER
                              C T A A A A C C C T T T G A G A A A A T C A A A C A A A G T
                                                2140                2150                2160
```

FIG.10K

```
LEU GLY GLN GLU LYS TYR ASP GLU ILE ASP
TTGGGGCAAGAAAAATACGACGAGATAGAC
        2170      2180      2190

ARG LEU GLY PHE ASN ALA TYR LYS ASP LEU
AGACTGGGCTTTAATGCTTATAAAGATTTA
        2200      2210      2220

ARG ASN GLU TRP ALA GLY TRP THR ASN ASP
CGCAACGAATGGGCGGGTTGGACTAATGAC
        2230      2240      2250

ASN SER GLN GLN ASN ALA ASN LYS GLY THR
AACAGCCAACAAAACGCCAATAAAGGCACG
        2260      2270      2280

ASP ASN ILE TYR GLN PRO ASN GLN ALA THR
GATAATATCTATCAGCCAAATCAAGCAACT
        2290      2300      2310

VAL VAL LYS ASP ASP LYS CYS LYS TYR SER
GTGGTCAAAGATGACAAATGTAAATATAGC
        2320      2330      2340

GLU THR ASN SER TYR ALA ASP CYS SER THR
GAGACCAACAGCTATGCTGATTGCTCAACC
        2350      2360      2370
```

FIG.10L

```
THR  ARG  HIS  ILE  SER  GLY  ASP  ASN  TYR  PHE
A C T C G C C A C A T C A G C G G T G A T A A T T A T T T C
          2380                2390                2400

ILE  ALA  LEU  LYS  ASP  ASN  MET  THR  ILE  ASN
A T C G C T T T A A A A G A C A A C A T G A C C A T C A A T
      2410                2420                2430

LYS  TYR  VAL  ASP  LEU  GLY  LEU  GLY  ALA  ARG
A A A T A T G T T G A T T T G G G C T G G G T G C T C G C
          2440                2450                2460

TYR  ASP  ARG  ILE  LYS  HIS  LYS  SER  ASP  VAL
T A T G A C A G A A T C A A A C A C A A A T C T G A T G T G
      2470                2480                2490

PRO  LEU  VAL  ASP  ASN  SER  ALA  SER  ASN  GLN
C C T T T G G T A G A C A A C A G T G C C A G C A A C C A G
          2500                2510                2520

LEU  SER  TRP  ASN  PHE  GLY  VAL  VAL  VAL  LYS
C T G T C T T G G A A T T T T G G C G T G G T C G T C A A G
      2530                2540                2550

PRO  THR  ASN  TRP  LEU  ASP  ILE  ALA  TYR  ARG
C C C A C C A A T T G G C T G G A C A T C G C T T A T A G A
          2560                2570                2580
```

09/142628

FIG.10M

```
SER   SER   GLN   GLY   PHE   ARG   MET   PRO   SER   PHE
A G C T C G C A A G G C T T T C G C A T G C C A A G T T T T
              2590                2600                2610
SER   GLU   MET   TYR   GLY   GLU   ARG   PHE   GLY   VAL
T C T G A A A T G T A T G G C G A A C G C T T T G G C G T A
              2620                2630                2640

THR   ILE   GLY   LYS   GLY   THR   GLN   HIS   GLY   CYS
A C C A T C G G T A A A G G C A C G C A A C A T G G C T G T
              2650                2660                2670
LYS   GLY   LEU   TYR   TYR   ILE   CYS   GLN   GLN   THR
A A G G G T C T T T A T T A C A T T T G T C A G C A G A C T
              2680                2690                2700

VAL   HIS   GLN   THR   LYS   LEU   LYS   PRO   GLU   LYS
G T C C A T C A A A C C A A G C T A A A A C C T G A A A A A
              2710                2720                2730
SER   PHE   ASN   GLN   GLU   ILE   GLY   ALA   THR   LEU
T C C T T T A A C C A A G A A A T C G G A G C G A C T T T A
              2740                2750                2760

HIS   ASN   HIS   LEU   GLY   SER   LEU   GLU   VAL   SER
C A T A A C C A C T T A G G C A G T C T T G A G G T T A G T
              2770                2780                2790
```

FIG.10N

```
                                    TYR PHE LYS ASN ARG TYR THR ASP LEU ILE
                                    T A T T T T A A A A A T C G C T A T A C C G A T T T G A T T
                                                2800                2810                2820
VAL GLY LYS SER GLU GLU ILE ARG THR LEU
G T T G G T A A A A G T G A A G A G A T T A G A A C C C T A
          2830                2840                2850
                                    THR GLN GLY ASP ASN ALA GLY LYS GLN ARG
                                    A C C C A A G G T G A T A A T G C A G G C A A A C A G C G T
                                                2860                2870                2880
GLY LYS GLY ASP LEU GLY PHE HIS ASN GLY
G G T A A A G G T G A T T T G G G C T T T C A T A A T G G G
          2890                2900                2910
                                    GLN ASP ALA ASP LEU THR GLY ILE ASN ILE
                                    C A A G A T G C T G A T T T G A C A G G C A T T A A C A T T
                                                2920                2930                2940
LEU GLY ARG LEU ASP LEU ASN ALA VAL ASN
C T T G G C A G A C T T G A C C T A A A C G C T G T C A A T
          2950                2960                2970
                                    SER ARG LEU PRO TYR GLY LEU TYR SER THR
                                    A G T C G C C T T C C C T A T G G A T T A T A C T C A A C A
                                                2980                2990                3000
```

09/142628

FIG.10O

```
LEU ALA TYR ASN LYS VAL ASP VAL LYS GLY
C T G G C T T A T A A C A A A G T T G A T G T T A A A G G A
            3010                3020                3030
                                  LYS THR LEU ASN PRO THR LEU ALA GLY THR
                                  A A A A C C T T A A A C C C A A C T T T G G C A G G A A C A
                                              3040                3050                3060
ASN ILE LEU PHE ASP ALA ILE GLN PRO SER
A A C A T A C T G T T T G A T G C C A T T C A G C C A T C T
            3070                3080                3090
                                  ARG TYR VAL VAL GLY LEU GLY TYR ASP ALA
                                  C G T T A T G T G G T G G G G C T T G G C T A T G A T G C C
                                              3100                3110                3120
PRO SER GLN LYS TRP GLY ALA ASN ALA ILE
C C A A G C C A A A A A T G G G G A G C A A A C G C C A T A
            3130                3140                3150
                                  PHE THR HIS SER ASP ALA LYS ASN PRO SER
                                  T T T A C C C A T T C T G A T G C C A A A A A T C C A A G C
                                              3160                3170                3180
GLU LEU LEU ALA ASP LYS ASN LEU GLY ASN
G A G C T T T T G G C A G A T A A G A A C T T A G G T A A T
            3190                3200                3210
```

09/142628

FIG.10P

```
GLY   ASN   ILE   GLN   THR   LYS   GLN   ALA   THR   LYS
G G C A A C A T T C A A A C A A A A C A A G C C A C C A A A
                  3220                      3230                      3240

ALA   LYS   SER   THR   PRO   TRP   GLN   THR   LEU   ASP
G C A A A A T C C A C G C C G T G G C A A A C A C T T G A T
                  3250                      3260                      3270

LEU   SER   GLY   TYR   VAL   ASN   ILE   LYS   ASP   ASN
T T G T C A G G T T A T G T A A A C A T A A A A G A T A A T
                  3280                      3290                      3300

PHE   THR   LEU   ARG   ALA   GLY   VAL   TYR   ASN   VAL
T T T A C C T T G C G T G C T G G C G T G T A C A A T G T A
                  3310                      3320                      3330

PHE   ASN   THR   TYR   TYR   THR   THR   TRP   GLU   ALA
T T T A A T A C C T A T T A C A C C A C T T G G G A G G C T
                  3340                      3350                      3360

LEU   ARG   GLN   THR   ALA   GLU   GLY   ALA   VAL   ASN
T T A C G C C A A A C A G C A G A A G G G G C G G T C A A T
                  3370                      3380                      3390

GLN   HIS   THR   GLY   LEU   SER   GLN   ASP   LYS   HIS
C A G C A T A C A G G A C T G A G C C A A G A T A A G C A T
                  3400                      3410                      3420
```

FIG.10Q

```
TYR GLY ARG TYR ALA ALA PRO GLY ARG ASN
T A T G G T C G C T A T G C C G C T C C T G G A C G C A A T
          3430                3440                3450
                                   TYR GLN LEU ALA LEU GLU MET LYS PHE ***
                                   T A C C A A T T G G C A C T T G A A A T G A A G T T T T A A
                                      S         3460                3470
3480
C C A G T G G C T T T G A T G T G A T C A T G C C A A A T C
          3490                3500                3510
                                   C C A A T C A A C C A A T G A A T A A A G C C C C C A T C T
                                             3520                3530                3540
A C C A T G A G G G C T T T A T T T T A T C A T C G C T G A
          3550                3560                3570
                                   G T A T G C T C T T A G C G G T C A T C A C T C A G A T T A
                                             3580                3590                3600
G T C A T T A A T T T A T T A G C G A T T A A T T T A T T A
          3610                3620                3630
                                   G T A A T C A C G C T G C T C T T T G A T G A T T T T A A G
                                             3640                3650                3660
```

FIG.11A

Q8 tbpB Sequence.

```
C C T A G G G C T G A C A G T A A C A A C A C T T T A T A C
                  10                  20                  30
                                                            A G C A C A T C A T T G A T T T A T T A C C C A A A T G C C
                                                                              40                  50                  60
A C A C G C T A T T A T C T T T T G G G G G C A G A C T T T
                  70                  80                  90
                                                            T A T G A T G A A A A A G T G C C A C A A G A C C C A T C T
                                                                             100                 110                 120
G A C A G C T A T G A G C G T C G T G G C A T A C G C A C A
                 130                 140                 150
                                                            G C T T G G G G G C A A G A A T G G G C G G G C G G T C T T
                                                                             160                 170                 180
T C A A G C C G T G C C C A A A T C A G C A T C A A C A A A
                 190                 200                 210
                                                            C G C C A T T A C C A A G G A G C A A A C C T A A C C A G C
                                                                             220                 230                 240
G G T G G A C A A A T T C G C C A G G A T A A A C A G A T G
                 250                 260                 270
                                                            C A A G C G T C T T T A T C G C T T T G G C A C A G A G A C
                                                                             280                 290                 300
```

FIG.11B

```
A T T C A C A A A T G G G G C A T C A C G C C A C G G C T G
                  310                 320                 330
                                          A C C A T C A G C A C A A A C A T C A A T A A A A G C A A T
                                                            340                 350                 360

G A C A T C A A G G C A A A T T A T C A C A A A A A T C A A
                  370                 380                 390
                                          A T G T T T G T T G A G T T T A G T C G C A T T T T T T G A
                                                            400                 410                 420

T G G G A T A A G C A T G C C C T A C T T T T G T T T T T T
                  430                 440                 450
                                          G T A A A A A A A T G T A C C A T C A T A G A C A A T A T C
                                                            460                 470                 480

A A G A A A A A A T C A A G A A A A A A G A T T A C A A A T
                  490                 500                 510
                                          T T A A T G A T A A T T G T T A T T G T T T A T G T T A T T
                                                            520                 530                 540

A T T T A T C A A T G T A A A T T T G C C G T A T T T T G T
                  550                 560                 570
                                          C C A T C A T A A A C G C A T T T A T C A A A T G C T C A A
                                                            580                 590                 600
```

FIG.11C

```
ATAAATACGCCAAATGCACATTGTCAACAT
         610       620       630
                    GCCAAAATAGGCATTAACAGACTTTTTTAG
                             640       650       660

ATAATACCATCAACCCATCAGAGGATTATT
         670       680       690
                      MET LYS HIS ILE PRO LEU THR THR LEU C
                    TTATGAAACACATTCCTTTAACCACACTGT
                             700       710       720

YS VAL ALA ILE SER ALA VAL LEU LEU THR
GTGTGGCAATCTCTGCCGTCTTATTAACCG
         730       740       750
                    ALA CYS GLY GLY SER SER GLY GLY PHE ASN P
                     CTTGTGGTGGTAGCAGTGGTGGTTTCAATC
                            760       770       780

RO PRO ALA SER THR PRO ILE PRO ASN ALA
CACCTGCCTCTACGCCCATCCCAAATGCAG
         790       800       810
                    GLY ASN SER GLY ASN ALA GLY ASN ALA GLY A
                     GTAATTCAGGTAATGCTGGCAATGCTGGCA
                            820       830       840
```

09/142628

FIG.11D

```
SN  ALA GLY GLY THR GLY GLY ALA ASN SER
ATGCTGGCGGTACTGGCGGTGCAAACTCTG
        850       860       870
                              GLY ALA GLY ASN ALA GLY GLY THR GLY GLY A
                              GTGCAGGTAATGCTGGCGGTACTGGCGGTG
                                      880       890       900

LA  ASN SER GLY ALA GLY SER ALA SER THR
CAAACTCTGGTGCAGGCAGTGCCAGCACAC
        910       920       930
                              PRO GLU PRO LYS TYR LYS ASP VAL PRO THR A
                              CAGAACCAAAATATAAAGATGTGCCAACCG
                                      940       950       960

SP  GLU ASN LYS LYS ALA GLU VAL SER GLY
ATGAAAATAAAAAAGCTGAAGTTTCAGGCA
        970       980       990
                              ILE GLN GLU PRO ALA MET GLY TYR GLY VAL G
                              TTCAAGAACCTGCCATGGGTTATGGCGTGG
                                     1000      1010      1020

LU  LEU LYS LEU ARG ASN TRP ILE PRO GLN
AATTAAAGCTTCGTAACTGGATACCACAAG
       1030      1040      1050
```

09/142628

FIG.11E

```
                                        GLU  GLN  GLU  GLU  HIS  ALA  LYS  ILE  ASN  THR  A
                                         A A C A G G A A G A A C A T G C C A A A A T C A A T A C A A
                                                    1060            1070            1080
SN  ASP  VAL  VAL  LYS  LEU  GLU  GLY  ASP  LEU
A T G A T G T T G T A A A A C T T G A A G G T G A C T T G A
            1090            1100            1110
                                        LYS  HIS  ASN  PRO  PHE  ASP  ASN  SER  ILE  TRP  G
                                         A G C A T A A T C C A T T T G A C A A C T C T A T T T G G C
                                                    1120            1130            1140
LN  ASN  ILE  LYS  ASN  SER  LYS  GLU  VAL  GLN
A A A A C A T C A A A A A T A G C A A A G A A G T A C A A A
            1150            1160            1170
                                        THR  VAL  TYR  ASN  GLN  GLU  LYS  GLN  ASN  ILE  G
                                         C T G T T T A C A A C C A A G A G A A G C A A A A C A T T G
                                                    1180            1190            1200
LU  ASP  GLN  ILE  LYS  ARG  GLU  ASN  LYS  GLN
A A G A T C A A A T C A A A A G A G A A A A T A A A C A A C
            1210            1220            1230
                                        ARG  PRO  ASP  LYS  LYS  LEU  ASP  ASP  VAL  ALA  L
                                         G C C C T G A C A A A A A A C T T G A T G A C G T G G C A C
                                                    1240            1250            1260
```

FIG.11F

```
EU  GLN  ALA  TYR  ILE  GLU  LYS  VAL  LEU  ASP
TACAAGCTTATATTGAAAAAGTTCTTGATG
          1270      1280      1290
                                ASP  ARG  LEU  THR  GLU  LEU  ALA  LYS  PRO  ILE  T
                                ACCGTCTAACAGAACTTGCTAAACCCATTT
                                          1300      1310      1320
YR  GLU  LYS  ASN  ILE  ASN  TYR  SER  HIS  ASP
ATGAAAAAAATATTAATTATTCACATGATA
          1330      1340      1350
                                LYS  GLN  ASN  LYS  ALA  ARG  THR  ARG  ASP  LEU  L
                                AGCAGAATAAAGCACGCACTCGTGATTTGA
                                          1360      1370      1380
YS  TYR  VAL  ARG  SER  GLY  TYR  ILE  TYR  ARG
AGTATGTGCGTTCTGGTTATATTTATCGCT
          1390      1400      1410
                                SER  GLY  TYR  SER  ASN  ILE  ILE  PRO  LYS  LYS  I
                                CAGGTTATTCTAATATCATTCCAAAGAAAA
                                          1420      1430      1440
LE  ALA  LYS  THR  GLY  PHE  ASP  GLY  ALA  LEU
TAGCTAAAACTGGTTTTGATGGTGCTTTAT
          1450      1460      1470
```

FIG.11G

```
PHE TYR GLN GLY THR GLN THR ALA LYS GLN L
TTTATCAAGGTACACAAACTGCTAAACAAT
        1480       1490       1500

EU PRO VAL SER GLN VAL LYS TYR LYS GLY
TGCCTGTATCTCAAGTTAAGTATAAAGGCA
        1510       1520       1530

THR TRP ASP PHE MET THR ASP ALA LYS LYS G
CTTGGGATTTTATGACCGATGCCAAAAAAG
        1540       1550       1560

LY GLN SER PHE SER SER PHE GLY THR SER
GACAATCATTTAGCAGTTTTGGTACATCGC
        1570       1580       1590

GLN ARG LEU ALA GLY ASP ARG TYR SER ALA M
AACGTCTTGCTGGTGATCGTTATAGTGCAA
        1600       1610       1620

ET SER TYR HIS GLU TYR PRO SER LEU LEU
TGTCTTACCATGAATACCCATCTTTATTAA
        1630       1640       1650

THR ASP GLU LYS ASN LYS PRO ASP ASN TYR A
CTGATGAGAAAAACAAACCAGATAATTATA
        1660       1670       1680
```

09/142028

FIG.11H

```
SN   GLY  GLU  TYR  GLY  HIS  SER  SER  GLU  PHE
A C G G T G A A T A T G G T C A T A G C A G T G A G T T T A
               1690                 1700                 1710
                    THR  VAL  ASP  PHE  SER  LYS  LYS  SER  LEU  LYS  G
                     C G G T A G A T T T T A G T A A A A A G A G C C T A A A A G
                                 1720                 1730                 1740
LY   GLU  LEU  SER  SER  ASN  ILE  GLN  ASP  GLY
G T G A G C T G T C T A G T A A C A T A C A A G A C G G C C
               1750                 1760                 1770
                    HIS  LYS  GLY  SER  VAL  ASN  LYS  THR  LYS  ARG  T
                     A T A A G G G C A G T G T T A A T A A A A C C A A A C G C T
                                 1780                 1790                 1800
YR   ASP  ILE  ASP  ALA  ASN  ILE  TYR  GLY  ASN
A T G A C A T C G A T G C C A A T A T C T A C G G C A A C C
               1810                 1820                 1830
                    ARG  PHE  ARG  GLY  SER  ALA  THR  ALA  SER  ASP  T
                     G C T T C C G T G G C A G T G C C A C C G C A A G C G A T A
                                 1840                 1850                 1860
HR   THR  GLU  ALA  SER  LYS  SER  LYS  HIS  PRO
C A A C A G A A G C A A G C A A A A G C A A A C A C C C C T
               1870                 1880                 1890
```

FIG.11.I

```
                                       PHE THR SER ASP ALA LYS ASN SER LEU GLU G
                                        T T A C C A G C G A T G C C A A A A A T A G C C T A G A A G
                                                      1900                1910                1920
LY  GLY PHE TYR GLY PRO ASN ALA GLU GLU
G C G G T T T T T A T G G A C C A A A C G C C G A G G A G C
           1930                1940                1950
                                       LEU ALA GLY LYS PHE LEU THR ASN ASP ASN L
                                        T G G C A G G T A A A T T C C T A A C C A A T G A C A A C A
                                                      1960                1970                1980
YS  LEU PHE GLY VAL PHE GLY ALA LYS ARG
A A C T C T T T G G C G T C T T T G G T G C T A A A C G A G
           1990                2000                2010
                                       GLU SER GLU ALA LYS GLU LYS THR GLU ALA I
                                        A G A G T G A A G C T A A G G A A A A A A C C G A A G C C A
                                                      2020                2030                2040
LE  LEU ASP ALA TYR ALA LEU GLY THR PHE
T C T T A G A T G C C T A T G C A C T T G G G A C A T T T A
           2050                2060                2070
                                       ASN LYS PRO GLY THR THR ASN PRO ALA PHE T
                                        A T A A A C C T G G T A C G A C C A A T C C C G C C T T T A
                                                      2080                2090                2100
```

09/142628

FIG.11J

```
HR  ALA  ASN  SER  LYS  LYS  GLU  LEU  ASP  ASN
C C G C T A A C A G C A A A A A A G A A C T G G A T A A C T
              2110                2120                2130
                                  PHE  GLY  ASN  ALA  LYS  LYS  LEU  VAL  LEU  GLY  S
                                  T T G G C A A T G C C A A A A A G T T G G T C T T G G G T T
                                                2140                2150                2160

ER  THR  VAL  ILE  ASP  LEU  VAL  PRO  THR  GLY
C T A C C G T C A T T G A T T T G G T G C C T A C C G G T G
              2170                2180                2190
                                  ALA  THR  LYS  ASP  VAL  ASN  GLU  PHE  LYS  GLU  L
                                  C C A C C A A A G A T G T C A A T G A A T T C A A A G A A A
                                                2200                2210                2220

YS  PRO  LYS  SER  ALA  THR  ASN  LYS  ALA  GLY
A G C C A A A G T C T G C C A C A A A C A A A G C G G G C G
              2230                2240                2250
                                  GLU  THR  LEU  MET  VAL  ASN  ASP  GLU  VAL  ILE  V
                                  A G A C T T T G A T G G T G A A T G A T G A A G T T A T C G
                                                2260                2270                2280

AL  LYS  THR  TYR  GLY  TYR  GLY  ARG  ASN  PHE
T C A A A A C C T A T G G C T A T G G C A G A A A C T T T G
              2290                2300                2310
```

FIG.11K

```
                                  GLU TYR LEU LYS PHE GLY GLU LEU SER ILE G
                                   A A T A C C T A A A A T T T G G T G A G C T T A G T A T C G
                                            2320                2330                2340
LY  GLY SER HIS SER VAL PHE LEU GLN GLY
G T G G T A G C C A T A G C G T C T T T T T A C A A G G C G
          2350                2360                2370
                                  GLU ARG THR ALA GLU LYS ALA VAL PRO THR G
                                   A A C G C A C C G C T G A G A A A G C C G T A C C A A C C G
                                            2380                2390                2400
LU  GLY THR ALA LYS TYR LEU GLY ASN TRP
A A G G C A C A G C C A A A T A T C T G G G G A A C T G G G
          2410                2420                2430
                                  VAL GLY TYR ILE THR GLY LYS ASP THR GLY T
                                   T A G G A T A C A T C A C A G G A A A G G A C A C A G G A A
                                            2440                2450                2460
HR  SER THR GLY LYS SER PHE ASN GLU ALA
C G A G C A C A G G A A A A A G C T T T A A T G A G G C C C
          2470                2480                2490
                                  GLN ASP ILE ALA ASP PHE ASP ILE ASP PHE G
                                   A A G A T A T T G C T G A T T T T G A C A T T G A C T T T G
                                            2500                2510                2520
```

FIG.11L

```
LU   ARG  LYS  SER  VAL  LYS  GLY  LYS  LEU  THR
A G A G A A A A T C A G T T A A A G G C A A A C T G A C C A
              2530                2540                2550
THR  GLN  GLY  ARG  GLN  ASP  PRO  VAL  PHE  ASN  I
C C C A A G G C C G C C A A G A C C C T G T A T T T A A C A
            2560                2570                2580

LE   THR  GLY  GLN  ILE  ALA  GLY  ASN  GLY  TRP
T C A C A G G T C A A A T C G C A G G T A A T G G C T G G A
              2590                2600                2610
THR  GLY  THR  ALA  SER  THR  ALA  LYS  ALA  ASN  V
C A G G C A C A G C C A G C A C C G C C A A A G C G A A C G
            2620                2630                2640

AL   GLY  GLY  TYR  LYS  ILE  ASP  SER  SER  SER
T A G G G G G C T A C A A G A T A G A T T C T A G C A G T A
              2650                2660                2670
THR  GLY  LYS  SER  ILE  VAL  ILE  GLU  ASN  ALA  L
C A G G C A A A T C C A T C G T C A T C G A A A A T G C C A
            2680                2690                2700

YS   VAL  THR  GLY  GLY  PHE  TYR  GLY  PRO  ASN
A G G T T A C A G G T G G C T T T T A T G G T C C A A A T G
              2710                2720                2730
```

FIG.11M

```
                              ALA ASN GLU MET GLY GLY SER PHE THR HIS A
                              C A A A C G A G A T G G G C G G G T C A T T T A C A C A C G
                                        2740                2750                2760
SP  THR ASP ASP SER LYS ALA SER VAL VAL
A T A C C G A T G A C A G T A A A G C C T C T G T G G T C T
          2770                2780                2790
                              PHE GLY THR LYS ARG GLN GLU GLU VAL LYS *
                              T T G G C A C A A A A A G A C A A G A A G A A G T T A A G T
                                        2800                2810                2820
**
A G T A A T T T A A A C A C A A T G C T T G G T T C G G C T
          2830                2840                2850
                              G A T G G G A T T G A C G C T T A A T C A A A C A T G A A T
                                        2860                2870                2880
G A T T A A G A T G A T A A A C C C A A G C C A T G C C A A
          2890                2900                2910
                              T G A T T G A T A G C A A C G A T G G C A G A T G A T G A G
                                        2920                2930                2940
T T T T C A T T A T C T G C C A T T A T T A T T G C T T A A
          2950                2960                2970
                              T T A T T G C T T G T C A T T T G G T G G T G T T A T C A C
                                        2980                2990                3000
```

FIG.11N

```
ATTAATCATTAAAATTAACATAATAAATGA
         3010      3020      3030
                              TTAAATGATATTTAATGAAAGTCAGGGTTA
                                       3040      3050      3060
TTTTGGTCATGGTTTTTCATGATTATTTAA
         3070      3080      3090
                              CTTATAATGCGTTATGGTTAGCAAAAAGCT
                                       3100      3110      3120
AAGTCTGTCAATGAAGCTATGGTGAGTGAT
         3130      3140      3150
                              TGTGCAAAAGATGGTCAAAAAAATCGGTAT
                                       3160      3170      3180
GGTGCTGTCAGGCGTGGTGATGGTTCTGTT
         3190      3200      3210
                              AATGATAATAACAACGCCAAGCCATGCTAC
                                       3220      3230      3240
TGCCAAGTTGTTGCCGACCTCTCAAGAAAA
         3250      3260      3270
                              TCCAACCAAAACTATGGTAGATAGCTTTGG
                                       3280      3290      3300
```

FIG.11.O

T C G T G A A A C G C C A C G A G G G G C A G T T C A G G G
3310 3320 3330

G C T A T T G C G T G C A A T T G C A G C A G A A G A C T A
3340 3350 3360

T G A G C T G G C T G C C A A C T A T T T G G A C G G C C G
3370 3380 3390

T T A T T T G G C A A A A A C C C A A A C G C C C A A T C G
3400 3410 3420

T G A G A T T G T T G A G C A
3430

FIG.12A

Tbp1 alignment

```
         10        20         30        40        50        60
MNQSKQNNKSKKSKQVLKLSALSLGLLNI--TQVALANTTADKAEA-TDKTNLVVVLDETVVT
.....---...................--..............-...............
.Q.QHLFR--------.NILC...--------MT..PVY----..NVQAEQAQEKQ..TIQ.K
.Q.QHLFR--------.NILC...--------MT..PAY----..NVQAGQAQEKQ..TIQ.K
.Q.QHLFR--------.NILC...--------MT..PAY----..NVQAGQAQEKQ..TIQ.K
.TKKPYFR-------LSIISC.LI.CYVKAE..SIKDTKE.ISS.VD.QS.E-DSE.ETIS..

                              70        80        90       100
                     AKKNA-RKANEVTGLGKVVKTAETINKEQVLNIRDLTRYDP      4223
                     .....-...................................      Q8
                     ...QKT.RD........L..SSD.LS...............      B16B6
                     ...QKT.RD........L....D.LS.....D.........      M982
                     ...QKT.RD........L....D.LS.....D.........      FA19
                     .E.IRD..D........II..S.S.SR..............      Eagan

       110       120       130       140       150        160
GIAVVEQGRGASSGYSIRGMDKNRVAVLVDGINQAQHYALQGPVAGKNYA-AGGAINEIEYEN
..................................................-............
.........................SLT...VS.I.S.TA.AALG.TRT.GSS..........
.........................SLT...LA.I.S.TA.AALG.TRT.GSS..........
.........................SLT...LA.I.S.TA.AALG.TRT.GSS..........
..S.................R....L....LP.T.S.VV.S.LVATSGYSGT..........
```

09/142628

FIG.12B

```
       170       180       190       200
VRSVEISKGANSSEYGSGALSGSVAFVTKTADDIIKDG     4223
......................................     Q8
.KA......S......N..A.....Q....A...GE.      B16B6
.KA......S..V.Q.....A.....Q......V.GE.     M982
.KA......S..V.Q.....A.....Q......V.GE.     FA19
.KA......GS.....N...A...T.QS.S.A..LEGD     Eagan

      210       220       230       240       250       260
KDWGVQTKTAYASKNNAWVNSVAAAGKAGSFSGLIIYTDRRGQEYKAHDDAYQGSQSFDRAVA
...............................................................
.Q..I.S....SG.DH.LTQ.L.L..RS.GAEA.L...K...R.IH..K..GK.V...N.L.L
RQ..I.S....SG..RGLTQ.I.L..RI.GAEA.L.H.G..AG.IR..E..GR.V...N.L.P
RQ..I.S....SG..RGLTQ.I.L..RI.GAEA.L.H.G.HAG.IR..EA.GR.V...N.LAP
.S..I...N..S...KGFTH.L.V...Q.G.E..A...Q.NSI.TQV.K..LK.V..Y..LI.

         270       280       290       300
TTD-----PNNRTFLIANECANFNYEACAAGGQTKLQAKPTN     4223
...-----...PK.............................     Q8
DE.KKEGGSQY.Y.IVEE..H..-.A..KNKL--.ED.SVKD     B16B6
VE.----SSEYAY.IVED..EGK...T.KSKP--.KDVVGKD     M982
VE.----GSKYAY.IVEE..K..GH.K.K.NP--.KDVVGED     FA19
...-----KSSGY.V.QG..P..DDK-...----.PP.TLST     Eagan
```

FIG.12C

```
     310       320       330        340        350       360
VRDKVNVKDYTGPNRLIPNPLTQDSKSLLLRPGYQLNDK-HYVGGVYEITKQNYAMQDKTVPA
......................................-.......................
E.KT.STQ....S...LA...EYG.Q.W.F...WH.DNR-....A.L.R.Q.TFDTR.M....
E.QT.STR.......FLAD..SYE.R.W.F...FRFENKR..I..IL.H.Q.TFDTR.M....
K.QT.STR.......FLAD..SYE.R.W.F...FRFENKR..I..IL.R.Q.TFDTR.M....
QSET.S.S....A..IK...MKYE.Q.WF..G..HFSEQ-..I..IF.F.Q.KFDIR.M.F..

     370       380          390       400
YLTVHDIEKSRLSNHAQA--NGYYQGNNLGERIRDTIGPD     4223
..............G..--...............A..AN     Q8
.F.SE.YVPGS.KGL------.K.S.D.KA..LFVQGEGS     B16B6
F..KAVFDANSKQAGSLPG-..K.A..HKYGGLFTNGENG     M982
F..KAVFDANQKQAGSLPG-..K.A..HKYGGLFTSGENN     FA19
..SPTERRDDSSRSFYPMQDH.A..HIE------------     Eagan

       410       420       430       440       450       460
----SGYGINYAHGVFYDEKHQKDELGLEYVYDSKGENKWFDDVRVSYDKQDITLRSQLTNTHC
----.............................................................
TLQGI...----T.....R.T.N.Y.V....HNADKDT.A.YA.L...R.G.D.DNR.QQ...
----ALV.AE.GT......T.T.S.Y.....TNADKDT.A.YA.L...R.G.G.DNHFQQ...
----APV.AE.GT......T.T.S.Y.....TNADKDT.A.YA.L...R.G.G.DNHFQQ...
----D.R.VK..S.LYF..H.R.Q.V.I..I.EN.NKAGII.KAVL.ANQ.N.I.D.YMRH...
```

09/142628

FIG.12D

```
                 470        480        490        500
STYPHIDKNCTPDVNKPFSVKEVDNNAYKEQHNLIKAVFN       4223
........................................       Q8
.HDGS-....R..G...Y.FYKS.RMI.E.SR..FQ...K       B16B6
.ADGS-..Y.R.SAD....YYKS.RVI.G.S.R.LQ.A.K       M982
.ADGS-..Y.R.SAD....YYKS.RVI.G.S.K.LQ.A.K       FA19
.L..NPS...R.TLD..Y.YYRS.R.V...K..MLQLNLE       Eagan

       510       520       530       540       550        560
KKMALGSTHHHINLQVGYDKFNSSLSREDYRLATHQSYQKLDYTPPSNPLPDKF-KPILGSNN
......N..............................................-.........
.AFDTAKIR.NLSINL...R.K.Q..HS..Y.QNAVQAYD.I-...KP.F.NGS-.------D
.SFDTAJUR.NKSVNK.F.R.S.B.RHQ..YYQHANRAYSSK-...KTAN.NGD-.------S
.SFDTAKIR.NLSVNL...T.G.N.RHQ..YYQSANRAYS.K-...Q.NGKKTS---PN.REK
..IQQNWLT.Q.VFNL.F.D.T.A.QHK..-.TRRVIATA-.SI.RK---.GETG..RN.LQS

                 570        580        590        600
KPICLDAYGYGHDHPQACNAKNSTYQNFAIKKGIEQYN         4223
R.....................................         Q8
N.YRVSIGK-----------------------------         B16B6
..YWVSIG.-----------------------------         M982
N.YWVSIGR-----------------------------         FA19
Q.YLYPKPEP----------------------------         Eagan
```

09/142628

FIG.12E

```
        610       620       630       640       650       660
QKTNTDKIDYQAIIDQYDKQNPNSTLKPFEKIKQSLGQEKYNKIDELGFKAYKDLRNEWAGWT
.-.........V........................DE..R...N.............
---------------------------------------------------------------
---------------------------------------------------------------
---------------------------------------------------------------
---------------------------------------------------------------

       670       680       690       700
NDNSQQNANKGRDNIYQPNQA-TVVKDDKCKYSETNS-Y      4223
.....................-.............-.        Q8
---------------------T..NTSPI.RFGN-.T-.      B16B6
---------------------GN..TGQI.LFGN-.T-.      M982
---------------------GN..TRQI.LFGN-.T-.      FA19
---------------------YFAGQDH-.N.QGSS.N.      Eagan

        710       720       730       740       750       760
ADCSTTRHISGDNYFIALKDNMTINKYVDLGLGARYDRIKHKSDVPLVDNSASNQLSWNFGVV
...............................................................
T..-.P.N.G.NG.YA.VQ..VRLGRWA.V.A.I...YRSTH.EDKS.STGTHRN....A...
T..-.P.S.N.KS.YA.VR..VRLGRWA.V.A.L...YRSTH..DGS.STGTHRT....A.I.
T..-.P.S.N.KS.YA.VR..VRLGRWA.V.A.L...YRSTH..DGS.STGTHRT....A.I.
R..-KV.L.K.K..YF.ARN..ALG........I...VSRT.ANESTISVGKFKNF...T.I.
```

09/142628

FIG.12F

```
        770       780       790       800
VKPTNWLDIAYRSSQGFRMPSFSEMYGERFGVTIGKG    4223
.....................................    Q8
L..FT.M.LT..A.T...L...A....W.A.ESLKTL    B16B6
L..AD...LT..T.T...L...A....W.S..OSKAV    M982
L..AD...LT..T.T...L...A....W.S.DK.KAV    FA19
I...E...LS..L.T...N........W.Y.GKNDEV    Eagan

      810       820       830       840       850       860
TQHGCKGLYYICQQTVHQTKLKPEKSFNQEIGATLHNHLGSLEVSYFKNRYTDLIVGKSEEIR
...............................................................
-------------------D........R.A.IVFKGDF.N..A...N.A.R...AFGY-.T.
-------------------.ID......K.A.IVFKGDF.N..A.W.N.A.R....RGY.AQI
-------------------.ID......K.A.IVFKGDF.N..A.W.N.A.R....RGY.AQI
----------------YVG.F...T.R...F.LA.KGDF.NI.I.H.S.A.RN..AFA-..LS

        870       880       890       900
TLTQGDNAGKQRGKGDLGFHNGQDADLTGINILGRLD    4223
......................N..........K..    Q8
.-----QN.QTSAS..P.YR.A.N.RIA......KI.    B16B6
K------N..EEA...PAYL.A.S.RI.......KI.    M982
K------D..EQV..NPAYL.A.S.RI.......KI.    FA19
K--------NGT...NY.Y..A.N.K.V.V..TAQ..    Eagan
```

FIG.12G

```
          910        920        930         940        950        960
LNAVNSRLPYGLYSTLAYNKVDVKGKTLNPTLAG-TNILFDAIQPSRYVVGLGYDAPSQKWGA
..............................-...............................
WHG.WGG..D.........RIK..DADIRADRTFV.SY....V......L.....H.DGI..I
WNG.WDK..E.W...F...R.H.RDIKKRADRTDIQSH................Q.EG...V
WNG.WDK..E.W...F...R.H.RDIKKRADRTDIQSH............S...Q.EG...V
F.GLWK.I...W.A.F...Q.K..DQKI.AG..SVSSY..........II.....H..NT..I
                                                      970       980        990       1000
                                                   NAIFTHSDAKNPSELLADKNLGNGNIQ-TKQATKAKSTP   4223
                                                   ...........................-...........   Q8
                                                   .TM..Y.K..SVD...GSQA.L...ANAK.A-ASRRTR.   B16B6
                                                   .GML.Y.K..EIT...GSRA.L...SRN..A-.ARRTR.   M982
                                                   .GML.Y.K..EIT...GSRA.L...SRN..A-.ARRTR.   FA19
                                                   .TM..Q.K..SQN...GKRA...-.SRDV.S-.RKLTRA   Eagan

      1010      1020      1030      1040      1050      1060      1070
WQTLDLSGYVNIKDNFTLRAGVYNVFNTYYTTWEALRQTAEGAVNQHTGLSQDKHYGRYAAPGRNYQLALEMKF*   4223
.........................................................................*   Q8
.YVT.V...Y...KHL........LL.YR.V...NV....G.:....---KNVGV.N.........TFS.....*   B16B6
.YIV.V...YT..KH.........LL.YR.V...NV....G......---KNVGV.N.........TFS.....*   M982
.YIV.V...YTV.KH.........LL.HR.V...NV....A......---KNVGV.N.........TFS.....*   FA19
.HI..V...YMANK.TM..L.I..L..YR.V....V....Q......---QNVGS.T....S....T.T.....*   Eagan
```

09/142628

FIG.13A

Tbp2 comparison

```
              10        20        30        40        50        60
MKHIPLTTLCVAISAV-LLTACGGS-GGSNPPAPTPIPNASGSGNTGNTGNAGGTDNT-ANAG
................-........S..F....S......GN...A..A......GGANSG..
.NN-..VNQAAMVLP.F..S..L.G-...-----------------------------------
.NN-..VNQAAMVLP.F..S..L.G-...-----------------------------------
.NN-..VNQAAMVLP.F..S..L.G-...-----------------------------------
..SV..ISGGLS----F..S..S.--...-----------------------------------

                        70        80        90        100
                NTGGT---NSGTGSANTPEPKYQDVPTEKNEKDK-VSSIQEPAM   4223
                .A...GGA...A...S......K....DE.K.AE-..G......   Q8
                -FDLDSVE---.VQDMHSK...EDEKS-QP.SQQD..ENSGA.-   B16B6
                -FDLDSVD---.EAPRPA-......SS..PQAQ.D-------QG   M982
                -FDLDSVD---.EAPRPA-......PSK.P.AR.D-------QG   FA19
                -FDVDNV--.N.P.---SK.R...DTSNQRK.S-NLKKLFI.SL   Eagan

       110       120       130       140       150
GYGMALSKINLHNRQDTPLD-EKNIITL--DGKKQVAEG-KKSPLPFS-LDV-ENKLLDGYIA
...VE.-.LRNWIP.EQEEH-A.IN-.N--.VV.LEGDL-.HN.FDN.IWQNIK.SKEVQTVY
-..F.V-.LPRR.AHFN.KYK..HKP.GSM.W--------..-LQRGEPNSFS.RDE.E----
...F.M-RLKRR.WYP--GAE.SEVK.NES.WEATGLPTKP.E-..KRQKS.I.KVET..D-S
...F.M-RFKRR.WHPSANPK.DEVK.KND.WEATGLPTEP.K-..LKQQS.ISEVETN.N-S
.G..K.VAQ..RGNKEPSFLN.DDY.-----------------SY..S.STI.KDVK.NNK-
```

09/142628

FIG.13B

```
160       170       180        190       200
KMNVADKNAIGDRIKKGNKEISDEELAKQIKEAVRKSHEFQQV-      4223
NQEKQNIEDQIK.EN.QRPDKKLDDV.L.AYIEKVLDDRLTELA      Q8
-------------K.R.SS.LI-.SKWEDGQSR.VGYTN.T---      B16B6
DIYSSPYLTPSNHQNG-----.AGNGVN.P.NQATGHEN..---      M982
..YTSPYLSQDADS-------.HANG.N.P.NE.TDYKK.----      FA19
--.G..--L..S-.---------..PSTTNPP.K----.HG.---     Eagan

      210       220       230       240       250       260
LSSLENKIFHSNDGTTKATTRDLKYVDYGY-YLANDGNYLTVKTDKLWNLGPVGGVFYNGTTT
KPIY.KN.NY.H.KQN..R.......RS..I.RSGYS.---IIPK.IAKT.FD.AL..Q..Q.
-----------------------..RS..V..-.KN.IDIKNNIV.F--..D.YLY.K.KEP
-----------------------..YS.WF.KH.ASEKDFSN.KI.S---.DD.YI..H.EK
----------------------...YS.WF.KH.KSEVKNENGLVSAKR-..D.YI..H.DK
----------------------...YS.LY.TPSWSLNDS-.N-.FY-..YY.YA..Y.NK.

      270       280       290        300
AKELPTQDAVKYKGHWDFMTDVANRRNRFSEVKENS--QA          4223
..Q..VSQ-.....T......-.KKGQS..SFGT-.QRL.          Q8
S....-SEKIT...T..YV..AME-KQ..-.GLG-.A..G          B16B6
PSRQ..ASGK.I...V.H.V..TKKGQD.R.IIQP.KK.G          M982
PSRQ..ASE..T...V.H.V..TKQGQK.NDIL.T.KG.G          FA19
.TN..VNGVA....T...I.ATK.-GK.YPLLSNG.H.--          Eagan
```

09/142628

FIG.13C

```
         310       320       330       340       350       360
GWYYGASSKD-EYNRLLTKEDSAPDGHSGEYGHSSEFTVNFKEKKLTGKLFSN---LQDRHKGN
.DR.S.M.YH-..PS...D.KNK..NYN...........D.SK.S.K.E.S..---I..G...S
DK-S..L.AL-.EGV.RNQAE-ASS..TD-F.MT...E.D.SD.TIK.T.YR.NRIT.NNSENK
DR.S.F.GDGS.EYSNKN-.STLK.D.E.-..FT.NLE.D.GN.......IR.NAS.NNNTNND
DK.S.F.GDEG.TTSNR.-DSNLN.K.E.-..FT.N.K.D.NN.......IR.NKVINTAASDG
--..RR-.AIP.DID.EN-DSKNG.-I.----LI...SADGGT.....Q.-.YTKRKTNNQPYE
         370       380       390       400

VTKTERYDIDANIHGNRFRGSATASNK--NDTSK-HPFTSDAN    4223
.N..K........Y............DTTEASK..-.......K    Q8
QI..T..T.Q.TL.....K.K.L.AD.--GA.NGS...I..SD    B16B6
KHT.QY.SL..Q.T....N.T...TD.K-ENET.L...V..SS    M982
Y.-..Y.SL..TLR....S.K.I.TD.PNTGGT.L...VF.SS    FA19
KK.L--.....D.YS.....TVKPTE.---.SEE-.....EGT    Eagan

         410       420       430            440
NRLEGGFYGPKGEELAGKFLTNDNKLFGVFGAKRESK------AEEKTE--------------
.S........NA.......................E------.K....--------------
S-..................S....VAA.....QKD.KDGENA.GPA..--------------
S-.S...F..Q....GFR..SD.Q.VAV.GS..TKD.LENGAA.SGS.G-AAASGGAAGTSSE
S-.S...F..Q....GFR..SD.G.VAV.GS..TKDST-----.NGNAP-AASSGPGAATMPS
--........NA...G....AT..RV....S..ETEETKKEALSK.TLIDGKLITFSTKKTDA
```

09/142628

FIG.13D

```
                                    450         460            470         480
                         -----AILDAYALGITFNISNAT--TFTPFTEKQLDNFGNAKKLV        4223
                         -----............KPGT.NPA..ANSK.E...........        Q8
                         -----TVI...RIT-----------GEEFKKE.I.S..DV...L        B16B6
                         NSKLTTV...VE.T-----------LNDKKI.N....S..AQ..        M982
                         ETRLTTV...VE.T-----------PDGKEI.N....S..TR..        FA19
                         KINATTSTA.NITTDITANTI.D--EKN.KTEDISS..E.DY.L        Eagan

      490                            510                          520
LGSTVIDLVP------------TDATK--NEFTKDK---PESATNEAGETLMVNDEVSV----
..........------------.G...DV...-.E.---.K....K...........I.----
VDGVELS.L.--SE-GNKAA--------------FQHEI.-----------------------
VDGIM.P.L.KDSESGNTQADKGKNGG--T...RKFEHT...DKKD.QAGTQINGAQTASNTA
VDGIM.P.L.--TESGNGQADKGKNGG--TD..YETTYT...DKKDTKAQTGAGGMQTASGTA
IDKYP.P.L.----------DKNIN---------------------------.FI.SK---
        530                        540

                         ---------KTYGKN------FEYLKFGELSIGGSH                 4223
                         ---------....YGRN----...............                 Q8
                         --QNGVKAT-------VCCSNLD.MS..K..KENKD                 B16B6
                         GDINGK--T...EVE-VCCSNLN...Y.M.TRKN.K                 M982
                         GVNGGQVGT...KVQ-VCCSNLN...Y.L..RENNN                 FA19
                         --HHTVGN-.R.KVEAVCCSNSDVKS.MYYEDPLKE                 Eagan
```

09/142628

FIG.13E

```
                                          550       560       570
--------------------------------------SVFLQGERTATTGEKAVPTTGTAKYLG
--------------------------------------..........---......E.......
--------------------------------------DM....V..PVSDVA.-R.EAN...R.
SAMQAGG-----------------NSSQADAKTEQVEQ.M.......---D..EI..DQNVV.R.
SVMQAVK-----------------NSSQADAKTKQIEQ.M.......---D.NKI.QEQGIV...
KETETETETEKDKEKEKEKDKDKEKQTAATTNTYYQ--..L.H..---PKDDI.K..S...H.

                         580       590       600       610
                   NWVGYIT-GKDTGTGTGKSFTDAQDVADFTIDFGNKSVSGK     4223
                   .......-......S.....NE...I...D...ER...K..     Q8
                   T.Y...AN.-TSWS.EA-.NQEGGNR.E.DV..ST.KI..T     B16B6
                   S.Y.H.AN.-TSWS.NA-.DKEGGNR.E..VN.AD.KIT..     M982
                   F.Y.R.AN.-TSWS.KA-.NATDGNR.K..VN.DR.EIT.T     FA19
                   S.Y....D..TSYSPS.DKKR.KNA..E.NV..AE.KLT.E     Eagan

      620       630       640       650       660       670
LITKGRQDPVFSITGQIAG--NGWTGTASTTKADAGGYKIDSSSTGKSIA--IKDANVTGGFYG
.T.Q.......N.......--.........A..NV.............V--.EN.K.......
.TA.D.TS.A.T..AM.KD--..FS.V.K.---GEN.FAL.PQN..N.HYTH.-E.T.S.....
.TAEN..AQT.T.E.M.Q.--..FE...K.---AES.FDL.QKN.TRTPKAY.T..K.K.....
.TAEN.SEAT.T.DAM.E.--..FK...K.---GND.FAP.QNNSTVTHKVH.AN.E.Q.....
.KRHDTGN.....EANFNNSS.AF....TA.-----NFV..GKNSQNKNTPINITTK.N.A...
```

09/142628

FIG.13F

```
         680                 690
PNANEMGGSFT-------------------NADDSKASV   4223
...........------------------HDT.......   Q8
K..I......SFPGNAPEGKQE--------------....   B16B6
.K.E.L..W.AYPGDKQTEKATATSSDG---.SAS.-.T.   M982
...E.L..W.AYPGNEQTKNATVESGNG---.SAS.-.T.   FA19
.K.S.L..Y..YNGNSTATNSESSSTVSSSS.SKNAP.A.   Eagan

     700
VFGTKRQQEV-K*   4223
.......E..-.*   Q8
...A....L.Q-*   B16B6
...A....P.Q-*   M982
...A...KL.-.*   FA19
...ARQ.V.TT.*   Eagan
```

Construction of TBP1 Expression Plasmid

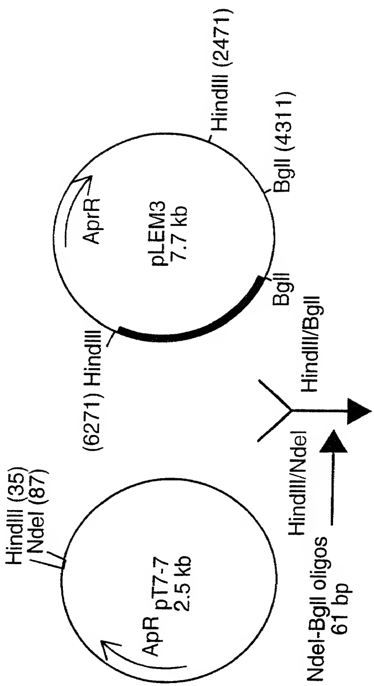


FIG.14A

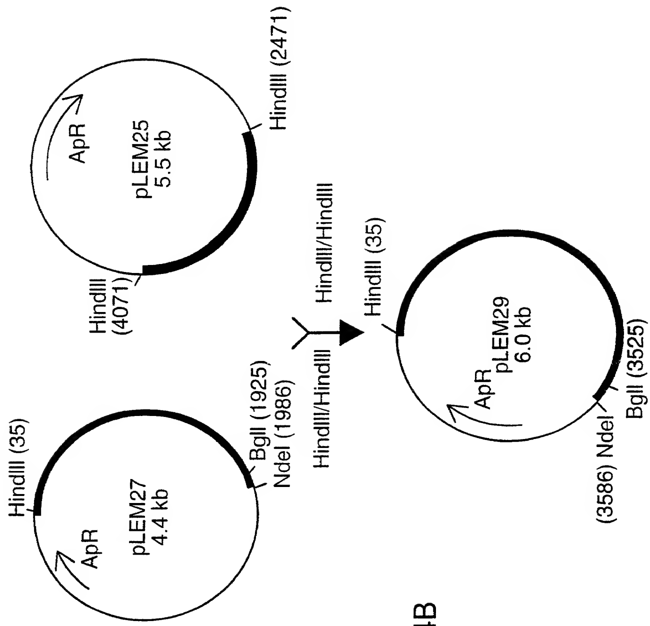
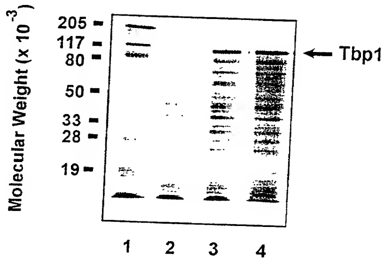

HindIII (2471)
ApR
pLEM25
5.5 kb
HindIII
(4071)
HindIII/HindIII
HindIII (35)
HindIII/HindIII
HindIII (35)
BglI (1925)
NdeI (1986)
ApR
pLEM27
4.4 kb
ApR
pLEM29
6.0 kb
BglI (3525)
(3586) NdeI

FIG.14B

Expression of rTbp1 in *E. coli*



1. Prestained molecular weight markers
2. pLEM29B-1 lysate, non-induced
3. pLEM29B-1 lysate, 1 hr post-induction
4. pLEM29B-1 lysate, 3 hr post-induction

Fig.15

Purification of Tbp1 from *E.Cole*

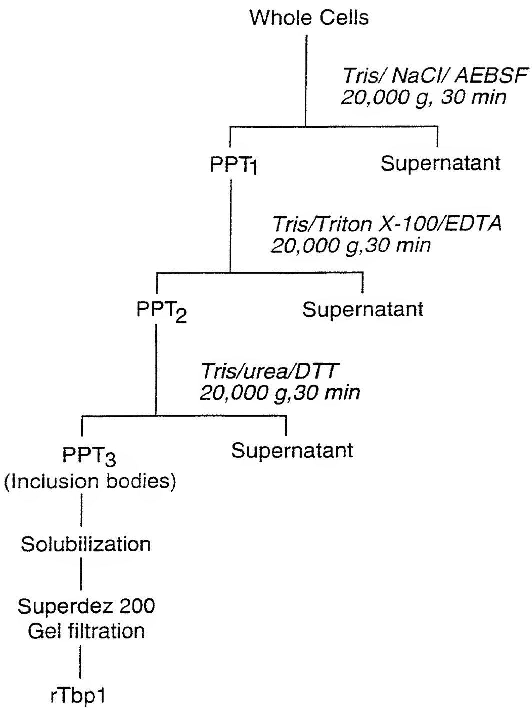
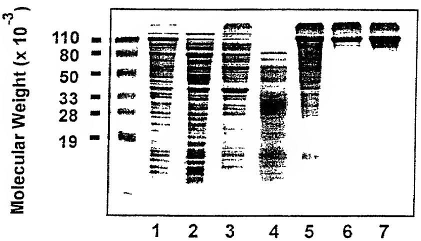


FIG.16

Purification of rTbp1 from *E. coli*



1. *E. coli* Whole cells

2. Soluble proteins after 50 mM Tris/ NaCl extraction

3. Soluble proteins after Tris/ Triton X-100/ EDTA extraction

4. Soluble proteins after Tris/ urea/ DTT extraction

5. Left-over pellet (rTbp1 inclusion bodies)

6.7. Purified rTbp1

Fig.17

CONSTRUCTION OF TBP2 EXPRESSION PLASMID

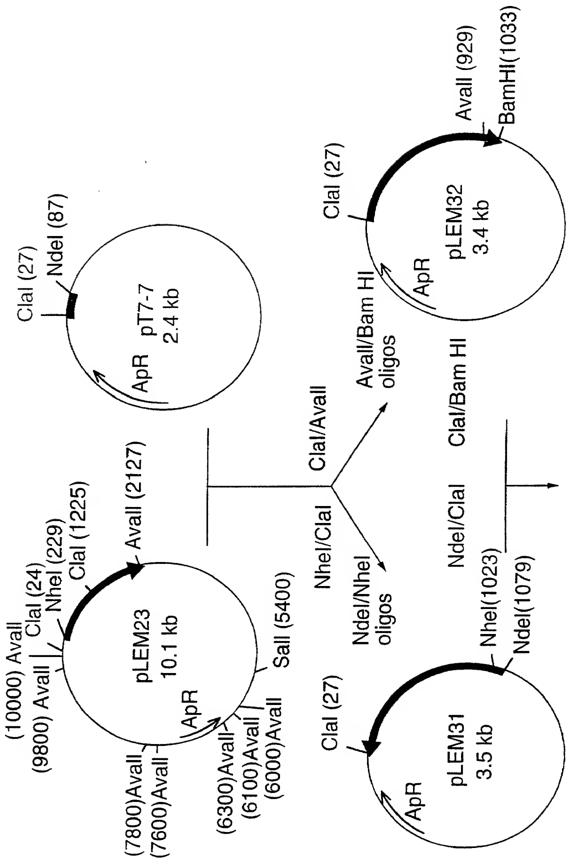
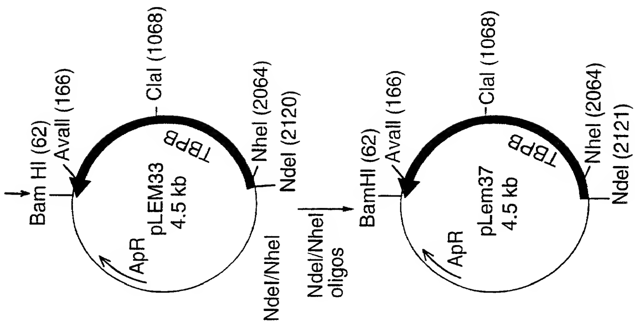
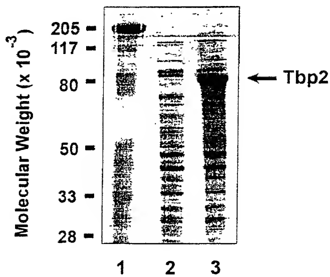


FIG.18A

FIG.18B


Bam HI (62)
AvaII (166)
ClaI (1068)
TBPB
NheI (2064)
NdeI (2120)
pLEM33
4.5 kb
ApR
NdeI/NheI
NdeI/NheI
oligos
BamHI (62)
AvaII (166)
ClaI (1068)
TBPB
NheI (2064)
NdeI (2121)
pLem37
4.5 kb
ApR

Expression of rTbp2 in *E. coli*



1. Prestained molecular weight markers
2. pLEM37B-2 lysate, non-induced
3. pLEM37B-2 lysate, induced

Fig.19

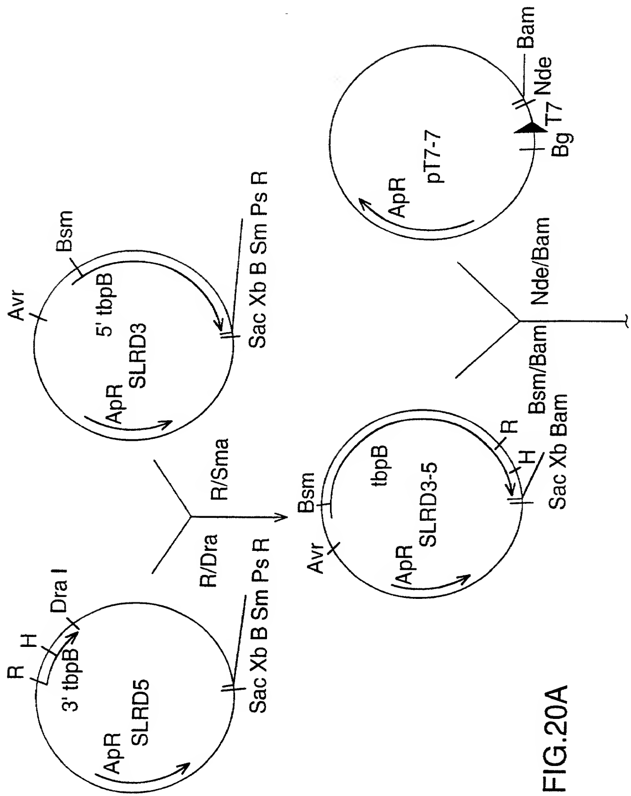

Bam
Nde
T7
Bg
pT7-7
ApR
Nde/Bam
Bsm/Bam
Bsm
Avr
5' tbpB
SLRD3
ApR
Sac Xb B Sm Ps R
R/Sma
R/Dra
tbpB
SLRD3-5
ApR
Bsm
Avr
R
H
Sac Xb Bam
R
H
Dra I
3' tbpB
SLRD5
ApR
Sac Xb B Sm Ps R

FIG.20A

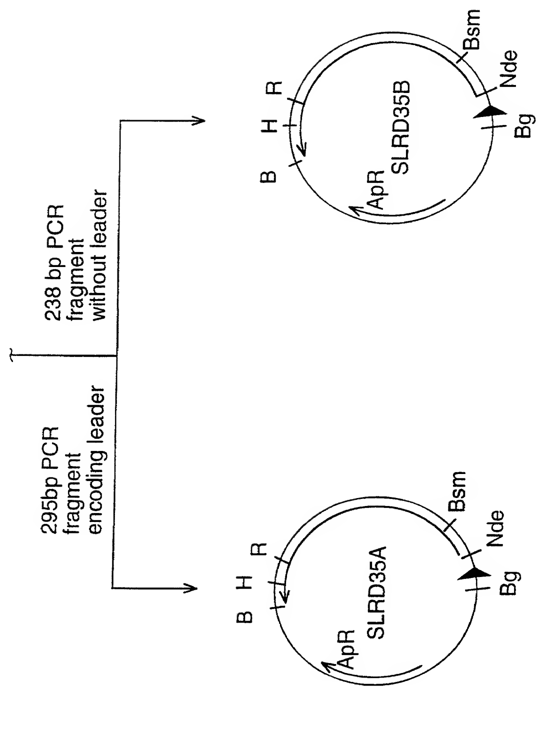
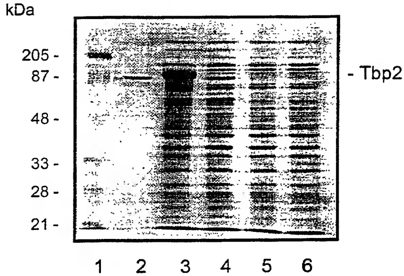

295bp PCR fragment encoding leader
238 bp PCR fragment without leader
SLRD35A
SLRD35B
B
H
R
Bsm
Nde
Bg
ApR

FIG.20B

Fig 21. Expression of Q8 rTbp2 protein in *E. coli*



1. Prestained molecular weight markers
2. 4223 rTbp2 protein
3. SLRD35A lysate, 3 hr post-induction
4. SLRD35B lysate, 3 hr post-induction
5. SLRD35A lysate, non-induced
6. SLRD35B lysate, non-induced

09/142628

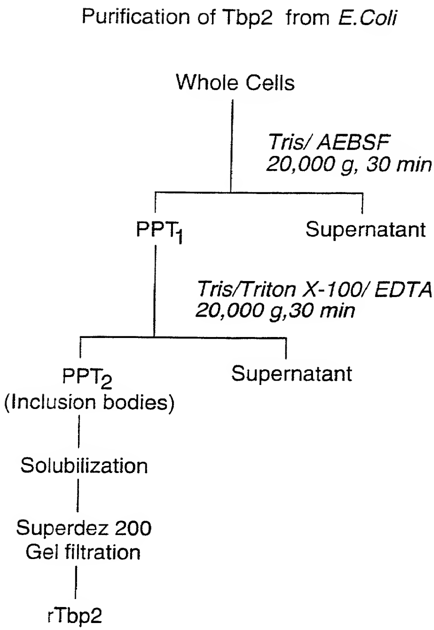
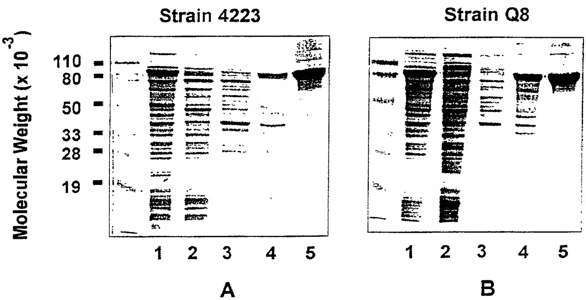


FIG.22

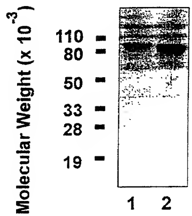
Purification of rTbp2 from *E. coli*



1. *E. coli* Whole cells
2. Soluble proteins after 50 mM Tris extraction
3. Soluble proteins after Tris/ Triton X-100/ EDTA extraction
4. Left-over pellet (rTbp2 inclusion bodies)
5. Purified rTbp2

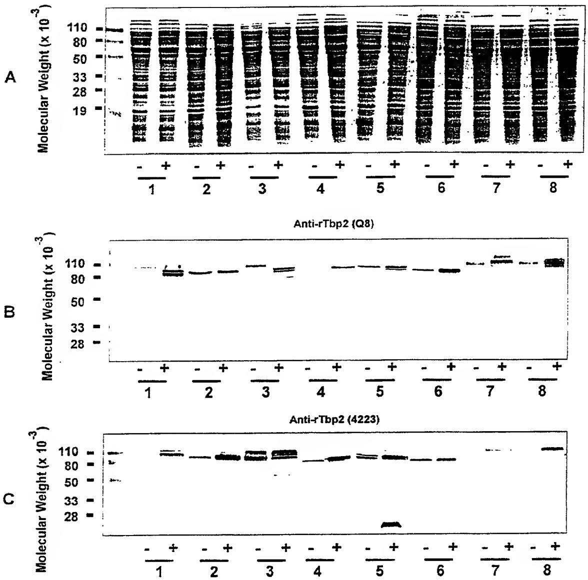
Fig.23

Binding of Tbp2 to Human Transferrin



1. rTbp2 (strain 4223)
2. rTbp2 (strain Q8)

Fig.24



1. Strain 5191
2. Strain 56
3. Strain 135
4. Strain 4223
5. Strain ATCC25240
6. Strain 585
7. Strain 3
8. Strain 8185

Fig.25

Figure 26 Restriction map of *M. catarrhalis* strain R1 *tbpB*

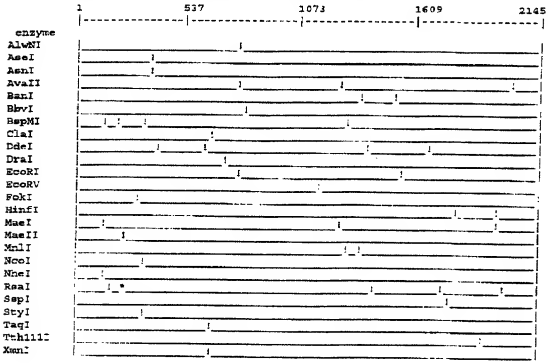


Figure 27 Nucleotide and deduced amino acid sequence of *M. catarrhalis* R1 *tbpB*

```
AAATTTGCCGTATTTTGTCTATCATAAATGCATTTATCATCAATGCCCAAACAAATACGCCAAATGCACAT

TGTCAGCATGCCAAAATAGGCATTAACAGACTTTTTTAGATAATACCATCAACCCATCAGAGGATTATTTT

                                          27                                     54
ATG AAA CAC ATT CCT TTA ACC ACA CTG TGT GTG GCA ATC TCT GCC GTC TTA TTA
MET Lys His Ile Pro Leu Thr Thr Leu Cys Val Ala Ile Ser Ala Val Leu Leu

                                          81                                    108
ACC GCT TGT GGT GGC AGT GGT GGT TCA AAT CCA CCT GCT CCT ACG CCC ATT CCA
Thr Ala Cys Gly Gly Ser Gly Gly Ser Asn Pro Pro Ala Pro Thr Pro Ile Pro

                                         135                                    162
AAT GCT AGC GGT TCA GGT AAT ACT GGC AAC ACT GGT AAT GCT GGC GGT ACT GAT
Asn Ala Ser Gly Ser Gly Asn Thr Gly Asn Thr Gly Asn Ala Gly Gly Thr Asp

                                         189                                    216
AAT ACA GCC AAT GCA GGT AAT ACA GGC GGT ACA AGC TCT GGT ACA GGC AGT GCC
Asn Thr Ala Asn Ala Gly Asn Thr Gly Gly Thr Ser Ser Gly Thr Gly Ser Ala

                                         243                                    270
AGC ACG TCA GAA CCA AAA TAT CAA GAT GTG CCA ACA ACG CCC AAT AAC AAA GAA
Ser Thr Ser Glu Pro Lys Tyr Gln Asp Val Pro Thr Thr Pro Asn Asn Lys Glu

                                         297                                    324
CAA GTT TCA TCC ATT CAA GAA CCT GCC ATG GGT TAT GGC ATG GCT TTG AGT AAA
Gln Val Ser Ser Ile Gln Glu Pro Ala MET Gly Tyr Gly MET Ala Leu Ser Lys

                                         351                                    378
ATT AAT CTA TAC GAC CAA CAA GAC ACG CCA TTA GAT GCA AAA AAT ATC ATT ACC
Ile Asn Leu Tyr Asp Gln Gln Asp Thr Pro Leu Asp Ala Lys Asn Ile Ile Thr

                                         405                                    432
TTA GAC GGT AAA AAA CAA GTT GCT GAC AAT CAA AAA TCA CCA TTG CCA TTT TCG
Leu Asp Gly Lys Lys Gln Val Ala Asp Asn Gln Lys Ser Pro Leu Pro Phe Ser

                                         459                                    486
TTA GAT GTA GAA AAT AAA TTG CTT GAT GGC TAT ATA GCA AAA ATG AAT GAA GCG
Leu Asp Val Glu Asn Lys Leu Leu Asp Gly Tyr Ile Ala Lys MET Asn Glu Ala

                                         513                                    540
GAT AAA AAT GCC ATT GGT GAA AGA ATT AAG AGA GAA AAT GAA CAA AAT AAA AAA
Asp Lys Asn Ala Ile Gly Glu Arg Ile Lys Arg Glu Asn Glu Gln Asn Lys Lys

                                         567                                    594
ATA TCC GAT GAA GAA CTT GCC AAA AAA ATC AAA GAA AAT GTG CGT AAA AGC CCT
Ile Ser Asp Glu Glu Leu Ala Lys Lys Ile Lys Glu Asn Val Arg Lys Ser Pro

                                         621                                    648
GAG TTT CAG CAA GTA TTA TCA TCG ATA AAA GCG AAA ACT TTC CAT TCA AAT GAC
Glu Phe Gln Gln Val Leu Ser Ser Ile Lys Ala Lys Thr Phe His Ser Asn Asp
```

FIG 27 (cont.)

```
                                          675                                       702
AAA ACA ACC AAA GCA ACC ACA CGA GAT TTA AAA TAT GTT GAT TAT GGT TAC TAC
Lys Thr Thr Lys Ala Thr Thr Arg Asp Leu Lys Tyr Val Asp Tyr Gly Tyr Tyr

                                          729                                       756
TTG GTG AAT GAT GCC AAT TAT CTA ACC GTC AAA ACA GAC AAC CCA AAA CTT TGG
Leu Val Asn Asp Ala Asn Tyr Leu Thr Val Lys Thr Asp Asn Pro Lys Leu Trp

                                          783                                       810
AAT TCA GGT CCT GTG GGC GGT GTG TTT TAT AAT GGC TCA ACG ACC GCC AAA GAG
Asn Ser Gly Pro Val Gly Gly Val Phe Tyr Asn Gly Ser Thr Thr Ala Lys Glu

                                          837                                       864
CTG CCC ACA CAA GAT GCG GTC AAA TAT AAA GGA CAT TGG GAC TTT ATG ACC GAT
Leu Pro Thr Gln Asp Ala Val Lys Tyr Lys Gly His Trp Asp Phe MET Thr Asp

                                          891                                       918
GTT GCC AAA AAA AGA AAC CGA TTT AGC GAA GTA AAA GAA ACC TAT CAA GCA GGC
Val Ala Lys Lys Arg Asn Arg Phe Ser Glu Val Lys Glu Thr Tyr Gln Ala Gly

                                          945                                       972
TGG TGG TAT GGG GCA TCT TCA AAA GAT GAA TAC AAC CGC TTA TTA ACC AAA GCA
Trp Trp Tyr Gly Ala Ser Ser Lys Asp Glu Tyr Asn Arg Leu Leu Thr Lys Ala

                                          999                                      1026
GAT GCC GCA CCT GAT AAT TAT AGC GGT GAA TAT GGT CAT AGC AGT GAA TTT ACT
Asp Ala Ala Pro Asp Asn Tyr Ser Gly Glu Tyr Gly His Ser Ser Glu Phe Thr

                                         1053                                      1080
GTT AAT TTT AAG GAA AAA AAA TTA ACA GGT GAG CTG TTT AGT AAC CTA CAA GAC
Val Asn Phe Lys Glu Lys Lys Leu Thr Gly Glu Leu Phe Ser Asn Leu Gln Asp

                                         1107                                      1134
AGC CAT AAA CAA AAA GTA ACC AAA ACA AAA CGC TAT GAT ATT AAG GCT GAT ATC
Ser His Lys Gln Lys Val Thr Lys Thr Lys Arg Tyr Asp Ile Lys Ala Asp Ile

                                         1161                                      1188
CAC GGC AAC CGC TTC CGT GGC AGT GCC ACC GCA AGC GAT AAG GCA GAA GAC AGC
His Gly Asn Arg Phe Arg Gly Ser Ala Thr Ala Ser Asp Lys Ala Glu Asp Ser

                                         1215                                      1242
AAA AGC AAA CAC CCC TTT ACC AGC GAT GCC AAA GAT AAG CTA GAA GGT GGT TTT
Lys Ser Lys His Pro Phe Thr Ser Asp Ala Lys Asp Lys Leu Glu Gly Gly Phe

                                         1269                                      1296
TAT GGA CCA AAA GGC GAG GAG CTG GCA GGT AAA TTC TTA ACC GAT GAT AAC AAA
Tyr Gly Pro Lys Gly Glu Glu Leu Ala Gly Lys Phe Leu Thr Asp Asp Asn Lys

                                         1323                                      1350
CTC TTT GGT GTC TTT GGT GCC AAA CAA GAG GGT AAT GTA GAA AAA ACC GAA GCC
Leu Phe Gly Val Phe Gly Ala Lys Gln Glu Gly Asn Val Glu Lys Thr Glu Ala
```

1377 1404
ATC TTA GAT GCT TAT GCA CTT GGG ACA TTT AAT AAA CCT GGT ACG ACC AAT CCC
Ile Leu Asp Ala Tyr Ala Leu Gly Thr Phe Asn Lys Pro Gly Thr Thr Asn Pro

1431 1458
GCC TTT ACC GCT AAC AGC AAA AAA GAA CTG GAT AAC TTT GGC AAT GCC AAA AAG
Ala Phe Thr Ala Asn Ser Lys Lys Glu Leu Asp Asn Phe Gly Asn Ala Lys Lys

1485 1512
TTG GTC TTG GGT TCT ACC GTC ATT GAT TTG GTG CCT ACT GAT GCC ACC AAA GAT
Leu Val Leu Gly Ser Thr Val Ile Asp Leu Val Pro Thr Asp Ala Thr Lys Asp

1539 1566
GTC AAT GAA TTC AAA GAA AAG CCA AAG TCT GCC ACA AAC AAA GCG GGC GAA ACT
Val Asn Glu Phe Lys Glu Lys Pro Lys Ser Ala Thr Asn Lys Ala Gly Glu Thr

1593 1620
TTG ATG GTG AAT GAT GAA GTT AGC GTC AAA ACC TAT GGC AAA AAC TTT GAA TAC
Leu MET Val Asn Asp Glu Val Ser Val Lys Thr Tyr Gly Lys Asn Phe Glu Tyr

1647 1674
CTA AAA TTT GGT GAG CTT AGT GTC GGT GGT AGC CAT AGC GTC TTT TTA CAA GGC
Leu Lys Phe Gly Glu Leu Ser Val Gly Gly Ser His Ser Val Phe Leu Gln Gly

1701 1728
GAA CGC ACC GCT ACC ACA GGC GAG AAA GCC GTA CCA ACC ACA GGC AAA GCC AAA
Glu Arg Thr Ala Thr Thr Gly Glu Lys Ala Val Pro Thr Thr Gly Lys Ala Lys

1755 1782
TAT TTG GGG AAC TGG GTA GGA TAT ATC ACA GGA GCG GAC TCA TCA AAA GGC TCT
Tyr Leu Gly Asn Trp Val Gly Tyr Ile Thr Gly Ala Asp Ser Ser Lys Gly Ser

1809 1836
ACC GAT GGC AAA GGC TTT ACC GAT GCC AAA GAT ATT GCT GAT TTT GAC ATT GAC
Thr Asp Gly Lys Gly Phe Thr Asp Ala Lys Asp Ile Ala Asp Phe Asp Ile Asp

1863 1890
TTT GAG AAA AAA TCA GTT AAT GGC AAA CTG ACC ACC AAA GAC CGC CAA GAC CCT
Phe Glu Lys Lys Ser Val Asn Gly Lys Leu Thr Thr Lys Asp Arg Gln Asp Pro

1917 1944
GTC TTT AAC ATC ACA GGT GAA ATC GCA GGC AAT GGC TGG ACA GGT AAA GCC AGC
Val Phe Asn Ile Thr Gly Glu Ile Ala Gly Asn Gly Trp Thr Gly Lys Ala Ser

1971 1998
ACC GCC GAA GCG AAC GCA GGG GGC TAT AAG ATA GAT TCT AGC AGT ACA GGC AAA
Thr Ala Glu Ala Asn Ala Gly Gly Tyr Lys Ile Asp Ser Ser Ser Thr Gly Lys

2025 2052
TCC ATC GTC ATC AAA GAT GCC GTG GTT ACA GGT GGC TTT TAT GGT CCA AAT GCA
Ser Ile Val Ile Lys Asp Ala Val Val Thr Gly Gly Phe Tyr Gly Pro Asn Ala

Fig 27 (cont)

								2079									2106
ACC	GAG	ATG	GGT	GGG	TCA	TTT	ACA	CAC	AAC	AGC	GGT	AAT	GAT	GGT	AAA	GTC	TCT
Thr	Glu	MET	Gly	Gly	Ser	Phe	Thr	His	Asn	Ser	Gly	Asn	Asp	Gly	Lys	Val	Ser

								2133				
GTG	GTC	TTT	GGC	ACA	AAA	AAA	CAA	GAA	GTT	AAG	AAG	TGA
Val	Val	Phe	Gly	Thr	Lys	Lys	Gln	Glu	Val	Lys	Lys	*

09/142628

FIG. 28

Alignment of *M. catarrhalis* Tbp2

```
          10        20        30        40        50        60        70        80        90       100
MKHIPLTTLCVAISAVLLTACGGSGGS-NPPAPTPIPNASGSGNTGNTGNAGGTDNTAN-AGNTGGT---NSGTGSANTPEPKYQDVPTEKNEKDKVSSIQEPAM   4223
.........................S.GF...S......GN...A..A......GGANSG...A...GGA...A...S......K....DE.K.AE..G......   Q8
.........................-..................................-...........---S......S.S.........TP.N.EQ.........   R1

         110       120       130       140       150       160       170       180       190       200
GYGMALSKINLHNRQDTPLDEKNIITLDGKKQVAEGKKSPLPFSLDVENKLLDGYIAKMNVADKNAIGDRIKKG---NKEISDEELAKQIK-EAVRKSHEFQQV   4223
...VE.KLR.WTPQEQEEHAKI.TNDVVKLEGDLKHNPFDNSIWQNIK.SKEVQTVYNQEKQNIEDQIK.EN.Q---RPDKKLDDV.L.AYI.K.LDDRLTELA   Q8
...........YDQ......A.............DNQ.......................E.......E...RENEQ..K........K..-.N....P.....   R1

         210       220       230       240       250       260       270       280       290       300
LSSLENKIFHSNDGTTKATTRDLKYVDYGYYLANDGNYLTVKTD--KLWNLGPVGGVFYNGTTTAKELPTQDAVKYKGHWDFMTDVANRRNRFSEVKENSQA   4223
KPIY.KN.NY.H.KQN..R.......RS..IYRSGYSNIIP.----.IAKT.FD.AL..Q..Q...Q..VSQ-.....T......AKKGQSFS.FGTSQRL.   Q8
...IKA.T.....K.................V..A........NP....S..........S.........................KK.........TY..   R1

         310       320       330       340       350       360       370       380       390       400
GWYYGASSKDEYNRLLTKEDSAPDGHSGEYGHSSEFTVNFKEKKLTGKLFSNLQDRHKGNVTKTERYDIDANIHGNRFRGSATASNKNDT--SKHPFTSDAN   4223
.DR.S.M.YH..PS...D.KNK..NYN...........D.SK.S.K.E.S..I..G...S.N..K........Y...........DTTEASK.........K   Q8
..W...............A.A...NY....................E.......S..QK....K....K.D.............D.AEDSK.........K   R1

         410       420       430       440       450       460       470       480       490       500
NRLEGGFYGPKGEELAGKFLTNDNKLFGVFGAKRESKAEEKTEAILDAYALGTFNTSNATT--FTPFTEKQLDNFGNAKKLVLGSTVIDLVPTDATK--NEFTK   4223
.S........NA.......................E.K................KPGT.NPA..ANSK.E......................G...DV...-.   Q8
DK..................D...........Q.GNV-................KPGT.NPA..ANSK.E.........................DV...-.   R1

         510       520       530       540       550       560       570       580       590       600
DKPESATNEAGETLMVNDEVSVKTYG--KNFEYLKFGELSIGGSHSVFLQCERTATTGEKAVPTTGTAKYLGNWVGYITGKDTGTGT--GKSFTDAQDVADFDI   4223
E..K....K...........I.....YGR...........................---......E...................S.--.....NE...I.....   Q8
E..K....K................--...............V........................K.............A.SSK.STD..G....K.I.....   R1

         610       620       630       640       650       660       670       680       690       700
DFGNKSVSGKLITKGRQDPVFSITGQIAGNGWTGTASTTKADAGGYKIDSSSTGKSIAIKDANVTGGFYGPNANEMGGSFTHNA-DDSKASVVFGTKRQQEVK-*   4223
..ER...K...T.Q.......N................A..NV..............V.EN.K...................DT-................E...-*   Q8
..EK...N...T..D......N...E........K...AE.N...............V....V..........T.........SGN.G.V.......K.-...K*   R1
```

1038-831 to -840

Express Mail Label No.

JUN 1 0 1999 PATENT & TRADEMARK OFFICE

Docket No. 1038-833 MIS

Declaration and Power of Attorney For Patent Application

English Language Declaration

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

TRANSFERRIN RECEPTOR GENES OF MORAXELLA

the specification of which

(check one)

☐ is attached hereto.

☒ was filed on March 7, 1997 as United States Application No. or PCT International Application Number PCT/CA97/00163

and was amended on ______________________

(if applicable)

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)			Priority Not Claimed
(Number)	(Country)	(Day/Month/Year Filed)	☐
(Number)	(Country)	(Day/Month/Year Filed)	☐
(Number)	(Country)	(Day/Month/Year Filed)	☐

Express ...

JUN 10 1999

Docket No.
1038-833 MIS

Declaration and Power of Attorney For Patent Application

English Language Declaration

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

TRANSFERRIN RECEPTOR GENES OF MORAXELLA

the specification of which

(check one)

☐ is attached hereto.

☒ was filed on March 7, 1997 as United States Application No. or PCT International Application Number PCT/CA97/00163

and was amended on ________________
(if applicable)

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)			Priority Not Claimed
(Number)	(Country)	(Day/Month/Year Filed)	☐
(Number)	(Country)	(Day/Month/Year Filed)	☐
(Number)	(Country)	(Day/Month/Year Filed)	☐

I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional application(s) listed below:

(Application Serial No.)	(Filing Date)

(Application Serial No.)	(Filing Date)

(Application Serial No.)	(Filing Date)

I hereby claim the benefit under 35 U. S. C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, C. F. R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
PCT/CA97/00163	07-March-1997	
08/778,570	03-January 1997	Pending
08/613,009	08-March 1996	Pending

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1038-831 to -840

Express Mail Label No.

OIPE JC17 JUN 1 0 1999 PATENT & TRADEMARK OFFICE

Docket No.
1038-833 MIS

Declaration and Power of Attorney For Patent Application

English Language Declaration

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

TRANSFERRIN RECEPTOR GENES OF MORAXELLA

the specification of which

(check one)

☐ is attached hereto.

☒ was filed on March 7, 1997 as United States Application No. or PCT International Application Number PCT/CA97/00163

and was amended on ______________________

(if applicable)

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)

			Priority Not Claimed
(Number)	(Country)	(Day/Month/Year Filed)	☐
(Number)	(Country)	(Day/Month/Year Filed)	☐
(Number)	(Country)	(Day/Month/Year Filed)	☐

I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional application(s) listed below:

(Application Serial No.)	(Filing Date)
(Application Serial No.)	(Filing Date)
(Application Serial No.)	(Filing Date)

I hereby claim the benefit under 35 U. S. C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, C. F. R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
PCT/CA97/00163	**07-March-1997**	
08/778,570	**03-January 1997**	**Pending**
08/613,009	**08-March 1996**	**Pending**

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. *(list name and registration number)*

Michael I. Stewart (24,973)

Send Correspondence to: **Sim & McBurney**
6th Floor, 330 University Avenue
Toronto, Ontario
Canada, M5G 1R7.

Direct Telephone Calls to: *(name and telephone number)*
(416) 595-1155

Full name of sole or first inventor **Lisa E. Myers**
Sole or first inventor's signature   Date
Residence **Guelph, Ontario, Canada**
Citizenship **Canadian**
Post Office Address **187 Elizabeth Street, Guelph, Ontario, Canada, N1E 2X5.**

Full name of second inventor, if any **Anthony B. Schryvers**
Second inventor's signature   Date
Residence **Calgary, Alberta, Canada**
Citizenship **Canadian**
Post Office Address **39 Edforth Road N.W., ,Calgary, Alberta, Canada, T3A 3V8.**

Full name of third inventor, if any **Robin E. Harkness**	
Third inventor's signature	Date
Residence **Willowdale, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **Apt 1706, 640 Sheppard Avenue East, Willowdale, Ontario, Canada, M2K 1B8.**	

Full name of fourth inventor, if any **Sheena M. Loosmore**	
Fourth inventor's signature	Date
Residence **Aurora, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **70 Crawford Rose Drive, Aurora, Ontario, Canada, L4G 4R4**	

Full name of fifth inventor, if any **Run-Pan Du**	
Fifth inventor's signature	Date
Residence **Thornhill, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **299 Chelwood Drive, Thorhnill, Ontario, Canada, L4J 7Y8.**	

Full name of sixth inventor, if any **Yan-Ping Yang**	
Sixth inventor's signature	Date
Residence **Willowdale, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **Apt. 709, 120 Torresdale Avenue, Willowdale, Ontario, Canada, M2R 3N7.**	

I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional application(s) listed below:

(Application Serial No.)	(Filing Date)
(Application Serial No.)	(Filing Date)
(Application Serial No.)	(Filing Date)

I hereby claim the benefit under 35 U. S. C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, C. F. R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
PCT/CA97/00163	**07-March-1997**	
08/778,570	**03-January 1997**	**Pending**
08/613,009	**08-March 1996**	**Pending**

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of seventh inventor, if any **Michel H. Klein**
Seventh inventor's signature   Date
Residence **Willowdale, Ontario, Canada.**
Citizenship **Canadian**
Post Office Address **16 Munro Boulevard, Willowdale, Ontario, Canada, M2P 1B9.**

Full name of eighth inventor, if
Eighth inventor's signature   Date
Residence
Citizenship
Post Office Address

Full name of ninth inventor, if any
Ninth inventor's signature   Date
Residence
Citizenship
Post Office Address

Full name of tenth inventor, if any
Tenth inventor's signature   Date
Residence
Citizenship
Post Office Address

Full name of third inventor, if any **Robin E. Harkness**	
Third inventor's signature	Date
Residence **Willowdale, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **Apt 1706, 640 Sheppard Avenue East, Willowdale, Ontario, Canada, M2K 1B8.**	

Full name of fourth inventor, if any **Sheena M. Loosmore**	
Fourth inventor's signature	Date
Residence **Aurora, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **70 Crawford Rose Drive, Aurora, Ontario, Canada, L4G 4R4.**	

Full name of fifth inventor, if any **Run-Pan Du**	
Fifth inventor's signature	Date
Residence **Thornhill, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **299 Chelwood Drive, Thorhnill, Ontario, Canada, L4J 7Y8.**	

Full name of sixth inventor, if any **Yan-Ping Yang**	
Sixth inventor's signature	Date
Residence **Willowdale, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **Apt. 709, 120 Torresdale Avenue, Willowdale, Ontario, Canada, M2R 3N7.**	

Full name of seventh inventor, if any
Michel H. Klein

Seventh inventor's signature | Date

Residence
Willowdale, Ontario, Canada.

Citizenship
Canadian

Post Office Address
16 Munro Boulevard, Willowdale, Ontario, Canada, M2P 1B9.

Full name of eighth inventor, if

Eighth inventor's signature | Date

Residence

Citizenship

Post Office Address

Full name of ninth inventor, if any

Ninth inventor's signature | Date

Residence

Citizenship

Post Office Address

Full name of tenth inventor, if any

Tenth inventor's signature | Date

Residence

Citizenship

Post Office Address

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. *(list name and registration number)*
Michael I. Stewart (24,973)

Send Correspondence to: **Sim & McBurney**
6th Floor, 330 University Avenue
Toronto, Ontario
Canada, M5G 1R7.

Direct Telephone Calls to: *(name and telephone number)*
(416) 595-1155

Full name of sole or first inventor **Lisa E. Myers**	
Sole or first inventor's signature Lisa M Myers	Date OCT 23 /98
Residence ROCKWOOD, ~~**Guelph,**~~ **Ontario, Canada**	CAX
Citizenship **Canadian**	
Post Office Address ~~**187 Elizabeth Street, Guelph, Ontario, Canada, N1E 2X5.**~~ 196 JACKSON ST, RR1 BOX 2, ROCKWOOD, ONTARIO N0B 2K0	

Full name of second inventor, if any **Anthony B. Schryvers**	
Second inventor's signature	Date
Residence **Calgary, Alberta, Canada**	
Citizenship **Canadian**	
Post Office Address **39 Edforth Road N.W., ,Calgary, Alberta, Canada, T3A 3V8.**	

Full name of third inventor, if any **Robin E. Harkness**	
Third inventor's signature	Date 23.10.98
Residence **Willowdale, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **Apt 1706, 640 Sheppard Avenue East, Willowdale, Ontario, Canada, M2K 1B8.**	

Full name of fourth inventor, if any **Sheena M. Loosmore**	
Fourth inventor's signature	Date
Residence **Aurora, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **70 Crawford Rose Drive, Aurora, Ontario, Canada, L4G 4R4.**	

Full name of fifth inventor, if any **Run-Pan Du**	
Fifth inventor's signature	oct.20.98 Date
Residence **Thornhill, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address ~~**299 Chelwood Drive, Thornhill, Ontario, Canada, L4J 7Y8.**~~ 88 TOWNSGATE Drive, Tornhill. Ontario, Canada, L4J 8A4	

Full name of sixth inventor, if any **Yan-Ping Yang**	
Sixth inventor's signature	Date 10-20-98
Residence **Willowdale, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **Apt. 709, 120 Torresdale Avenue, Willowdale, Ontario, Canada, M2R 3N7.**	

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. *(list name and registration number)*

Michael I. Stewart (24,973)

Send Correspondence to: **Sim & McBurney**
6th Floor, 330 University Avenue
Toronto, Ontario
Canada, M5G 1R7.

Direct Telephone Calls to: *(name and telephone number)*
(416) 595-1155

Full name of sole or first inventor **Lisa E. Myers**	
Sole or first inventor's signature	Date
Residence **Guelph, Ontario, Canada**	
Citizenship **Canadian**	
Post Office Address **187 Elizabeth Street, Guelph, Ontario, Canada, N1E 2X5.**	

Full name of second inventor, if any **Anthony B. Schryvers**	
Second inventor's signature Anthony B. S.	Date October 2, 1998
Residence **Calgary, Alberta, Canada**	
Citizenship **Canadian**	
Post Office Address **39 Edforth Road N.W., ,Calgary, Alberta, Canada, T3A 3V8.**	

Full name of seventh inventor, if any **Michel H. Klein**
Seventh inventor's signature [signature] Date October 5, 1998
Residence **Willowdale, Ontario, Canada.**
Citizenship **Canadian**
Post Office Address **16 Munro Boulevard, Willowdale, Ontario, Canada, M2P 1B9.**

Full name of eighth inventor, if
Eighth inventor's signature Date
Residence
Citizenship
Post Office Address

Full name of ninth inventor, if any
Ninth inventor's signature Date
Residence
Citizenship
Post Office Address

Full name of tenth inventor, if any
Tenth inventor's signature Date
Residence
Citizenship
Post Office Address